



UNIVERSIDAD DE JAÉN
Escuela Politécnica Superior de Linares

Trabajo Fin de Grado

EVALUACIÓN DEL USO DE TÉCNICAS NMF PARA IDENTIFICACIÓN DE INSECTOS

Alumno: David Abad Cruz

Tutor: Prof. D. Damián Martínez Muñoz

Depto.: Ingeniería de Telecomunicación

Junio, 2017

UNIVERSIDAD DE JAÉN
Escuela Politécnica Superior de Linares

Trabajo Fin de Grado

EVALUACIÓN DEL USO DE TÉCNICAS NMF PARA IDENTIFICACIÓN DE INSECTOS

Alumno: David Abad Cruz

Tutor: Prof. D. Damián Martínez Muñoz

Depto.: Ingeniería de Telecomunicación

Junio, 2017

ÍNDICE GENERAL

INTRODUCCIÓN	12
1.1 CONTEXTO	12
1.2 BÚSQUEDA DE INFORMACIÓN E INVESTIGACIÓN	12
1.3 OBJETIVOS	13
MATERIALES Y MÉTODOS	14
2.1 ESTADO DEL ARTE	14
2.1.1 NMF (Non-negative Matrix Factorization)	14
2.1.2 Aplicaciones de NMF	16
2.1.2.1 Mejora de la detección de voz eliminando ruido de fondo [1]	16
2.1.2.2 Análisis de datos de microarrays [2]	18
2.1.2.3 Separación de fuentes habladas grabadas en un sólo canal [3]	19
2.1.2.4 Transcripción de música polifónica [4]	20
2.1.2.5 Fusión de imágenes pancromáticas y multiespectrales capturadas mediante satélite [6]	20
2.1.3 Detección de insectos mediante captación de señales	21
2.1.3.1 Captura y procesamiento de sonidos [18]	21
2.1.3.2 Registro de vibraciones producidas por insectos (I) [13]	22
2.1.3.3 Registro de vibraciones producidas por insectos (II) [13]	22
2.2 BASE DE DATOS DE INSECTOS UTILIZADA	22
2.2.1 Cicadae (Cigarra)	24
2.2.2 Acrididae (Saltamontes)	25
2.2.3 Gryllidae (Grillo)	25
2.3 BASE DE DATOS DE RUIDO UTILIZADA	26
2.4 MODELO UTILIZADO	29
2.4.1 Principios básicos de la separación de fuentes con NMF [1]	30
2.4.1.1 Uso de NMF para la separación de fuentes	30
2.4.2 Descomposición del sonido del insecto y separación del ruido [1]	31
2.4.2.1 Modelo de señal	32
2.4.3 Descomposición del algoritmo [1]	32
2.4.3.1 Restricciones de ruido suave	33
2.4.3.2 Restricciones de ruido impulsivo	33
2.4.3.3 Restricciones de ruido de armónico	34
2.4.3.4 Estimación de parámetros	35

2.4.3.5 Síntesis de la señal	35
RESULTADOS Y DISCUSIÓN.....	36
3.1 ADAPTACIÓN DEL SOFTWARE A LAS NECESIDADES DEL TFG	36
3.2 ACTUALIZACIÓN DEL SOFTWARE	37
3.3 RESULTADOS OBTENIDOS.....	37
3.3.1 <i>Resultados obtenidos añadiendo ruido blanco a los ficheros de sonidos de insectos</i>	37
3.3.2 <i>Resultados añadiendo Noizeus para diferentes valores de los elementos de entrada del software</i>	41
3.3.2.1 Ruido aeropuerto.....	43
3.3.2.2 Ruido de conversación	45
3.3.2.3 Ruido de coche.....	48
3.3.2.4 Ruido de sala de exposiciones.....	50
3.3.2.5 Ruido de restaurante	52
3.3.2.6 Ruido de estación de tren.....	54
3.3.2.7 Ruido de calle.....	56
3.3.2.8 Ruido de tren	58
3.3.2.9 Primeras conclusiones sobre el uso de diferentes bases en el reconocimiento de insectos	60
3.3.3 <i>Resultados añadiendo Noizeus con 5 bases de caracterización</i>	60
3.3.3.1 Conclusiones añadiendo Noizeus con número fijo de bases y variación de valores de sparseness y smoothness.....	68
3.3.4 <i>Resultados añadiendo Noizeus manteniendo valores de sparseness y smoothness y modificando las bases de caracterización.....</i>	68
3.3.4.1 Conclusiones añadiendo Noizeus con valores fijos de sparseness y smoothness y variación de número de bases de caracterización	76
3.3.5 <i>Estudio aislado de insectos 13 y 17</i>	76
3.3.5.1 Conclusiones del estudio de insectos 13, 17 y 22.....	83
CONCLUSIONES.....	84
REFERENCIAS BIBLIOGRÁFICAS	87

ÍNDICE DE FIGURAS

FIGURA 1 MICRÓFONO AVISOFT-ULTRASOUND GATE 116	23
FIGURA 2 ÓRGANO ESTRIDULADOR CICADAЕ	24
FIGURA 3 ÓRGANO ESTRIDULADOR ACRIDIDAE(I)	25
FIGURA 4 ÓRGANO ESTRIDULADOR ACRIDIDAE (II).....	25
FIGURA 5 GRYLLIDAE.....	26
FIGURA 6 RESPUESTA EN FRECUENCIA FILTRO IRS	26
FIGURA 7 SEÑAL RAMPA POSITIVA DE 10 MUESTRAS.....	27
FIGURA 8 MUESTRAS FINALES DEL FICHERO 1 DE RUIDO AIRPORT MULTIPLICADO POR SEÑAL RAMPA NEGATIVA	28
FIGURA 9 MUESTRAS INICIALES DEL FICHERO 1 DE RUIDO AIRPORT MULTIPLICADO POR SEÑAL RAMPA POSITIVA.....	28
FIGURA 10 UNIÓN FICHEROS 1 Y 2 DE RUIDO AIRPORT	29
FIGURA 11 PORCENTAJE DE ÉXITO PARA DIFERENTES NIVELES DE SNR DE TODOS LOS INSECTOS AÑADIENDO RUIDO BLANCO GAUSSIANO. EL TOTAL DE OPERACIONES POR CADA NIVEL HA SIDO DE 161. LA CONFIGURACIÓN USADA HA SIDO 5 BASES DE CARACTERIZACIÓN PARA EL INSECTO Y LOS DIFERENTES TIPOS DE RUIDO CON UN VALOR PARA SPARSENESS Y SMOOTHNESS IGUAL A 1.	38
FIGURA 12 INSECTO 1 CON RUIDO BLANCO GAUSSIANO Y UN NIVEL DE -15 dB DE SNR. EN ROJO, EL INSECTO ORIGINAL. EN AZUL, EL INSECTO CON EL RUIDO AÑADIDO.....	39
FIGURA 13 INSECTO 13	39
FIGURA 14 ESPECTROGRAMA INSECTO 13	40
FIGURA 15 INSECTO 13 CON RUIDO BLANCO Y UN NIVEL DE SNR DE 0 dB	40
FIGURA 16 ESPECTROGRAMA INSECTO 13 CON RUIDO BLANCO Y UN NIVEL DE SNR DE 0 dB	41
FIGURA 17 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE AEROPUERTO. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 5 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	44
FIGURA 18 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE AEROPUERTO. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 3 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	44
FIGURA 19 RUIDO DE AEROPUERTO	45
FIGURA 20 ESPECTROGRAMA RUIDO DE AEROPUERTO.....	45
FIGURA 21 RUIDO DE CONVERSACIÓN	46
FIGURA 22 ESPECTROGRAMA RUIDO DE CONVERSACIÓN	46

FIGURA 23 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE CONVERSACIÓN. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 5 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	47
FIGURA 24 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE CONVERSACIÓN. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 3 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	47
FIGURA 25 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE COCHE. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 5 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	48
FIGURA 26 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE COCHE. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 3 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	49
FIGURA 27 RUIDO DE COCHE	49
FIGURA 28 ESPECTROGRAMA RUIDO DE COCHE.....	50
FIGURA 29 RUIDO DE SALA DE EXPOSICIONES	50
FIGURA 30 ESPECTROGRAMA RUIDO DE SALA DE EXPOSICIONES.....	51
FIGURA 31 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE SALA DE EXPOSICIONES. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 5 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.	51
FIGURA 32 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE SALA DE EXPOSICIONES. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 3 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.	52
FIGURA 33 RUIDO DE RESTAURANTE	52
FIGURA 34 ESPECTROGRAMA RUIDO DE RESTAURANTE	53
FIGURA 35 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE RESTAURANTE. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 5 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	53
FIGURA 36 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE RESTAURANTE. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 3	

BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	54
FIGURA 37 RUIDO DE ESTACIÓN	54
FIGURA 38 ESPECTROGRAMA RUIDO DE ESTACIÓN DE TREN	55
FIGURA 39 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE ESTACIÓN DE TREN. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 5 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	55
FIGURA 40 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE ESTACIÓN DE TREN. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 3 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	56
FIGURA 41 RUIDO DE CALLE.....	56
FIGURA 42 ESPECTROGRAMA RUIDO DE CALLE.....	57
FIGURA 43 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE CALLE. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 5 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	57
FIGURA 44 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE CALLE. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 3 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	58
FIGURA 45 RUIDO DE TREN	58
FIGURA 46 ESPECTROGRAMA RUIDO TRAIN.....	59
FIGURA 47 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE TREN. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 5 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	59
FIGURA 48 PORCENTAJE DE ÉXITO EN DETECCIÓN PARA TODOS LOS INSECTOS CON RUIDO DE TREN. 1150 OPERACIONES POR CADA NIVEL DE SNR. CONFIGURACIÓN DE 3 BASES DE CARACTERIZACIÓN DE RUIDO Y DE INSECTO Y CON VALOR 1 PARA SPARSENESS Y SMOOTHNESS.....	60
FIGURA 49 GRÁFICO DE PORCENTAJE DE ÉXITO EN DETECCIÓN CON 5 BASES PARA TODOS LOS TIPOS DE RUIDO Y CARACTERIZACIÓN DEL INSECTO. LOS VALORES MOSTRADOS CORRESPONDEN LA COMBINACIÓN DE SPARSENESS Y SMOOTHNESS RESPECTIVAMENTE. SE TOMAN EN CUENTA TODOS LOS TIPOS DE RUIDO Y LOS VALORES DE SNR (20 HASTA - 10DB). EL NÚMERO TOTAL DE OPERACIONES POR CADA COMBINACIÓN ES 1288.....	61

FIGURA 50 GRÁFICO DE PORCENTAJE DE ÉXITO EN DETECCIÓN CON 5 BASES PARA TODOS LOS TIPOS DE RUIDO Y CARACTERIZACIÓN DEL INSECTO. LOS VALORES MOSTRADOS AL ÉXITO EN LA DETECCIÓN DE INSECTOS PARA CADA NIVEL DE NSR. SE TOMAN EN CUENTA TODOS LOS TIPOS DE RUIDO. EL NÚMERO TOTAL DE OPERACIONES POR CADA COMBINACIÓN ES 1056.....	62
FIGURA 51 INSECTO 12	63
FIGURA 52 ESPECTROGRAMA INSECTO 12	63
FIGURA 53 DETECCIÓN EXITOSA PARA EL INSECTO 12 CON TODOS LOS TIPOS DE RUIDO Y PARA TODOS LOS VALORES DE SNR. EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE KSP Y KSM ES 336.	64
FIGURA 54 INSECTO 12 CON RUIDO DE CALLE CON UN NIVEL DE SNR DE 0 dB.....	64
FIGURA 55 ESPECTROGRAMA INSECTO 12 CON RUIDO DE CALLE CON NIVEL DE SNR DE 0 DB	65
FIGURA 56 INSECTO 14	65
FIGURA 57 ESPECTROGRAMA INSECTO 14	66
FIGURA 58 PORCENTAJE DE ÉXITO EN DETECCIÓN DEL INSECTO 14 PARA TODOS LOS RUIDOS Y TODOS LOS VALORES DE SNR. EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE KSP Y KSM ES 336.	66
FIGURA 59 INSECTO 14 CON RUIDO DE TREN Y UN NIVEL DE SNR DE 0 DB.....	67
FIGURA 60 ESPECTROGRAMA INSECTO 14 CON RUIDO DE TREN 0 DB	67
FIGURA 61 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 1. SE TOMAN EN CUENTA TODOS LOS TIPOS DE RUIDO Y LOS VALORES DE SNR (20 HASTA -20dB). EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 1656.....	69
FIGURA 62 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NIVELES DE SNR DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 1. EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 368.....	70
FIGURA 63 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 3. SE TOMAN EN CUENTA TODOS LOS TIPOS DE RUIDO Y LOS VALORES DE SNR (20 HASTA -20dB). EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 1656.....	71
FIGURA 64 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NIVELES DE SNR DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 3. EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 368.....	71
FIGURA 65 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 10. SE TOMAN EN CUENTA TODOS LOS TIPOS DE RUIDO Y LOS VALORES DE SNR (20 HASTA -10dB). EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 1288.....	72

FIGURA 66 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NIVELES DE SNR DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 10. EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 368.	72
FIGURA 67 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 30. SE TOMAN EN CUENTA TODOS LOS TIPOS DE RUIDO Y LOS VALORES DE SNR (20 HASTA -10dB). EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 1288.	73
FIGURA 68 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NIVELES DE SNR DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 30. EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 368.	73
FIGURA 69 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 100. SE TOMAN EN CUENTA TODOS LOS TIPOS DE RUIDO Y LOS VALORES DE SNR (20 HASTA -10dB). EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 1288.	74
FIGURA 70 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NIVELES DE SNR DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 100. EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 368.	74
FIGURA 71 PORCENTAJE DE ÉXITO EN LA IDENTIFICACIÓN DE INSECTOS PARA DIFERENTES NÚMEROS DE BASES CON KSP Y KSM FIJO A 1. SE TOMAN EN CUENTA TODOS LOS TIPOS DE RUIDO Y LOS VALORES DE SNR (20 HASTA -20dB). EL NÚMERO TOTAL DE OPERACIONES POR CADA VALOR DE BASES ES 1656.	75
FIGURA 72 INSECTO 13	76
FIGURA 73 ESPECTROGRAMA INSECTO 13	77
FIGURA 74 INSECTO 17	77
FIGURA 75 ESPECTROGRAMA INSECTO 17	77
FIGURA 76 INSECTO 22	78
FIGURA 77 ESPECTROGRAMA INSECTO 22	78
FIGURA 78 INSECTO 17 CON RUIDO DE CONVERSACIÓN Y UN NIVEL DE SNR DE 0 DB.....	79
FIGURA 79 ESPECTROGRAMA INSECTO 17 CON RUIDO DE CONVERSACIÓN Y UN NIVEL DE SNR DE 0 DB	79
FIGURA 80 INSECTO 13 CON RUIDO DE CALLE CON UN NIVEL DE SNR DE 10 DB.....	80
FIGURA 81 ESPECTROGRAMA INSECTO 13 CON RUIDO DE CALLE PARA UN NIVEL DE SNR DE 10 DB.....	80
FIGURA 82 GRÁFICO DE PORCENTAJE DE ÉXITO EN EL RECONOCIMIENTO DE INSECTOS CON UNA CONFIGURACIÓN PARA LOS INSECTOS Y LOS TIPOS DE RUIDO Y UN VALOR 1 PARA SPARSENESS Y SMOOTHNESS. EL TOTAL DE OPERACIONES POR CADA NIVEL DE SNR HAN SIDO 168.....	81

FIGURA 83 GRÁFICO DE PORCENTAJE DE ÉXITO EN EL RECONOCIMIENTO DE INSECTOS CON UNA CONFIGURACIÓN PARA LOS INSECTOS Y LOS TIPOS DE RUIDO Y UN VALOR 10 PARA SPARSENESS Y SMOOTHNESS. EL TOTAL DE OPERACIONES POR CADA NIVEL DE SNR HAN SIDO 168.....	81
FIGURA 84 GRÁFICO DE PORCENTAJE DE ÉXITO EN EL RECONOCIMIENTO DE INSECTOS CON UNA CONFIGURACIÓN PARA LOS INSECTOS Y LOS TIPOS DE RUIDO Y UN VALOR 100 PARA SPARSENESS Y SMOOTHNESS. EL TOTAL DE OPERACIONES POR CADA NIVEL DE SNR HAN SIDO 168.....	82
FIGURA 85 GRÁFICO DE PORCENTAJE DE ÉXITO EN EL RECONOCIMIENTO DE INSECTOS CON UNA CONFIGURACIÓN DE 1 BASE DE CARACTERIZACIÓN PARA LOS INSECTOS Y LOS TIPOS DE RUIDO Y UN VALOR 1 PARA SPARSENESS Y SMOOTHNESS. EL TOTAL DE OPERACIONES POR CADA NIVEL DE SNR HAN SIDO 264.....	82
FIGURA 86 GRÁFICO DE PORCENTAJE DE ÉXITO EN EL RECONOCIMIENTO DE INSECTOS CON UNA CONFIGURACIÓN DE 3 BASES DE CARACTERIZACIÓN PARA LOS INSECTOS Y LOS TIPOS DE RUIDO Y UN VALOR 1 PARA SPARSENESS Y SMOOTHNESS. EL TOTAL DE OPERACIONES POR CADA NIVEL DE SNR HAN SIDO 2024.....	84
FIGURA 87 GRÁFICO DE PORCENTAJE DE ÉXITO EN EL RECONOCIMIENTO DE INSECTOS CON UNA CONFIGURACIÓN DE 5 BASES DE CARACTERIZACIÓN PARA LOS INSECTOS Y LOS TIPOS DE RUIDO Y UN VALOR 1 PARA SPARSENESS Y SMOOTHNESS. EL TOTAL DE OPERACIONES POR CADA NIVEL DE SNR HAN SIDO 2024.....	85
FIGURA 88 GRÁFICO DE PORCENTAJE DE ÉXITO EN EL RECONOCIMIENTO DE INSECTOS CON UNA CONFIGURACIÓN DE 1 BASE DE CARACTERIZACIÓN PARA LOS INSECTOS Y LOS TIPOS DE RUIDO Y UN VALOR 1 PARA SPARSENESS Y SMOOTHNESS. EL TOTAL DE OPERACIONES POR CADA NIVEL DE SNR HAN SIDO 2024.....	85

ÍNDICE DE TABLAS

TABLA 1 BASE DE DATOS DE INSECTOS	24
TABLA 2 PARÁMETROS DE ENTRADA DEL SOFTWARE.....	42
TABLA 3 ERRORES EN IDENTIFICACIÓN INSECTO 2 CON RUIDO DE AEROPUERTO Y NIVEL DE SNR DE 10 DB.....	43

INTRODUCCIÓN

1.1 Contexto

Actualmente, se pueden utilizar diferentes técnicas o métodos para conocer, dentro de un área determinada, la biodiversidad existente. Por ejemplo, para conocer el tipo de animales que habita en un determinado lugar, nos bastaría con una búsqueda visual a nuestro alrededor y podríamos sacar unas primeras conclusiones de la biodiversidad existente en la zona.

Sin embargo, si nos queremos centrar en los insectos, todo se vuelve más complicado al ser seres muy pequeños y casi imperceptibles, en el caso de que estén bien camuflados con el entorno. Por tanto, para poder llevar a cabo un estudio sobre ellos, es necesario utilizar técnicas como la captura para, posteriormente, llegar a la conclusión del tipo de insectos existente en la zona, tras analizar los resultados obtenidos. Suele ser especialmente complicado y costoso por el hecho de tener tiempos de espera largos y, en el caso del uso trampas, se deben usar diferentes por cada especie que se quiera estudiar. En nuestro caso, nos vamos a centrar en un método de identificación basado en la captura de sonidos y posterior procesamiento, mediante diferentes algoritmos, para conseguir reconocer el insecto en cuestión.

1.2 Búsqueda de información e investigación

Para la realización de este Trabajo Fin de Grado, en adelante TFG, se ha realizado una búsqueda sobre el funcionamiento, implementación y aplicaciones del algoritmo NMF para el correcto uso de las diferentes variables que pueda contener para su evaluación.

A continuación, se ha indagado en el uso técnicas más frecuentes utilizadas para la identificación de insectos y se ha buscado información sobre la morfología de los insectos estudiados para conocer, con unas pequeñas pinceladas, el órgano u órganos capaces de producir el sonido que los caracteriza y en qué fase de su ciclo de vida se produce.

Otro aspecto a tener en cuenta, ha sido la búsqueda de una base de datos de sonidos de insectos suficiente y una base de datos de ruidos con los que mezclar los anteriores para proceder a realizar todas las pruebas necesarias.

1.3 Objetivos

Es objeto de este Trabajo Fin de Grado, la evaluación del uso de técnicas NMF (Non-negative Matrix Factorization) para la identificación de insectos mediante un algoritmo, por el cual, se pueda identificar el tipo de insecto en diferentes entornos y valorar si es posible la implementación del sistema en dicho campo.

Se pretende definir una matriz que contenga las bases que caractericen el sonido de cada insecto y tras analizar la matriz de activación de las anteriores, con grabaciones en otros ambientes, se tratará de identificar los insectos que los generan.

Se partirá de una base de datos en los que tendremos sonidos limpios de diferentes insectos y se procederá a introducirles diferentes tipos de ruido para observar y estudiar hasta qué punto es viable esta técnica.

Se hará un estudio de las diferentes configuraciones posibles de implementación para lograr optimizar el sistema en el que sea capaz de reconocer los insectos en mayor medida.

MATERIALES Y MÉTODOS

2.1 Estado del arte

En esta sección, se hará un repaso de la técnica de factorización de matrices no negativas (NMF) y se comentarán algunas aplicaciones de ésta. Se nombrarán otros sistemas de detección de insectos y su funcionamiento. Por último se explicará cómo producen el sonido algunos de los insectos utilizados y la procedencia de los ficheros utilizados al igual que los ficheros de ruido utilizados para las diferentes pruebas llevadas a cabo en este TFG.

2.1.1 NMF (*Non-negative Matrix Factorization*)

La factorización no negativa de la matriz, fue propuesta por primera vez en 1999 por Lee y Seung [7] inspirados por el uso de la factorización de la matriz positiva [8].

La factorización de matrices no negativas, desde este momento NMF, parte de la no negatividad de las muestras ofrecidas por las fuentes que se van a estudiar. Existen fuentes cuya información son no negativas, y como consecuencia, sólo tienen sentido físico cuando se cumple esta no negatividad. En el caso de reconocimiento de patrones, muchos de los problemas que nos encontramos trata de tener en cuenta las restricciones de no negatividad. Un ejemplo son las imágenes digitales donde los píxeles que la componen no son negativos [2]. Se ha comprobado que el uso de NMF es útil para la descomposición de datos multivariados.

NMF se aplica a algoritmos de aprendizaje no supervisado, como el análisis de componentes principales y la cuantificación vectorial. Se pueden entender estos algoritmos como factorizar una matriz de datos sujeta a diferentes restricciones y dependiendo de éstas, se tienen propiedades de representación muy diferentes. El análisis de componentes principales sólo cumple una mínima parte de restricciones de ortogonalidad, teniendo como resultado una representación muy distribuida que se utiliza para cancelaciones para generar variabilidad. La cuantificación vectorial por su parte, utiliza una restricción mucho más dura de *winnertake-all* que resulta en una agrupación de datos mutuamente excluyentes [16].

Dada una matriz no negativa V , se tienen que encontrar los factores de la matriz no negativa W y H .

$$V \approx WH \quad (1)$$

NMF se puede aplicar al análisis estadístico de datos multivariados de la siguiente manera: dado un conjunto de vectores de datos dimensionales multivariados, se posicionan en columnas en una matriz V $n \times m$ donde m es el número de ejemplo en el conjunto de datos. Entonces esta matriz es factorizada en otra matriz W $n \times r$ y otra matriz H $r \times m$. Donde r suele ser más pequeño que n o m , para que las matrices W y H sean más pequeñas que la matriz original V . Ésta última es la matriz original y queda comprimida en las anteriores.

La aproximación de la ecuación (1) se puede escribir columnas por columnas como $v \approx Wh$, donde v y h son las columnas correspondientes a V y H . Cada vector v es aproximado por una combinación lineal de las columnas de W sopesadas por las componentes de h . Por tanto, W se puede considerar que contiene unas bases que optimizan la aproximación lineal de los datos en V . Por ello, sólo se puede tener una buena aproximación si los vectores contenidos en W descubren la estructura latente en los datos.

El coste de las factorizaciones viene definida por unas funciones que cuantifican la calidad de la aproximación. La función de coste se construye usando la medida de distancia entre dos matrices no negativas A y B . Una medida útil es el cuadrado de la distancia euclidiana entre A y B .

$$\|A - B\|^2 = \sum_{ij} (A_{ij} - B_{ij})^2 \quad (2)$$

Otra medida es:

$$D(A||B) = \sum_{ij} \left(A_{ij} \log \frac{A_{ij}}{B_{ij}} - A_{ij} + B_{ij} \right) \quad (3)$$

En las dos fórmulas anteriores, está limitado por cero y desaparece si $A = B$. En esta última (3), no se puede llamar “distancia” porque no es simétrica en A y B y nos referimos a ella como “divergencia” de A y B . Se reduce a entropía relativa cuando $\sum_{ij} A_{ij} = \sum_{ij} B_{ij} = 1$, por lo que A y B son consideradas distribuciones de probabilidad normalizada.

Se aplican unas reglas de actualización multiplicativos de W y H para tener un buen compromiso entre velocidad y facilidad de implementación. Para $\|V - WH\|$ tenemos las siguientes reglas:

$$H_{a\mu} \leftarrow H_{a\mu} \frac{(W^T V)_{a\mu}}{(W^T W H)_{a\mu}} \quad (4)$$

$$W_{ia} \leftarrow W_{ia} \frac{(V H^T)_{ia}}{(W (H H^T))_{ia}} \quad (5)$$

Para la divergencia $D(V \| WH)$ tenemos las siguientes reglas:

$$H_{a\mu} \leftarrow H_{a\mu} \frac{\sum_i W_{ia} V_{i\mu} (W - V)_{i\mu}}{\sum_k W_{ka}} \quad (6)$$

$$W_{ia} \leftarrow W_{ia} \frac{\sum_\mu W_{a\mu} V_{i\mu} (W - V)_{i\mu}}{\sum_v W_{av}} \quad (7)$$

2.1.2 Aplicaciones de NMF

Se va a explicar algunos casos donde se aplica NMF, y comprobaremos que son ámbitos diferentes, como transcripción polifónica o estudios de medicina.

2.1.2.1 Mejora de la detección de voz eliminando ruido de fondo [1]

Es la base de éste TFG, ya que se ha utilizado un software similar pero con diversas modificaciones para ser útil con el sonido generado por los insectos.

Consiste en la creación de un pequeño diccionario de posibles fonemas, obtenido de diferentes grabaciones libres de ruido, que se procesan para caracterizar cada uno de ellos. El diccionario estará compuesto de diferentes bases para cada fonema.

Realizado este primer paso, se procederá al reconocimiento propiamente dicho. Se parte de grabaciones con un discurso con interferencias y ruido, para realizar una separación de la señal de audio en 4 componentes.

La primera de ellas contendrá la información de la voz, para poder realizar el reconocimiento sobre ésta sin que tenga interferencias de ningún tipo. Las 3 componentes restantes, contendrán la información de ruido existente que afectaba a la grabación. Se separará en varios tipos de ruido: ruido blanco, ruido percusivo y ruido armónico.

Para relacionar la componente de voz con los diferentes posibles fonemas, se activarán en la fase de reconocimiento, ciertas bases de nuestro diccionario indicándonos qué tipo de fonema es el pronunciado en cada caso.

La separación del habla con ruido en segundo plano y otras interferencias acústicas es una de las líneas de investigación más populares en la línea de investigación. Su aplicación incluye muchas situaciones donde la señal del habla está contaminada y debe de ser restaurada a la original. El problema viene para mezclas de un solo canal, donde no hay información espacial para separar las diferentes fuentes. En este caso, se aplica el método comentado anteriormente, donde se pretende descomponer el sonido con interferencias en las 4 componentes para intentar recomponer la señal.

Para distinguir los sonidos vocales se realiza una estimación aproximada del comportamiento en la producción del sonido vocal según la fisionomía humana. Se establece un modelo fuente/filtro para el reconocimiento de la voz. Para ello se realiza una descomposición partiendo del vector s_t :

$$s_t = \sum_{j=1}^J h_{jt}^s w_j^s \quad (8)$$

en una parte de excitación s_t^{F0} multiplicado por una parte de filtro s_t^Φ , que es compuesto por una combinación lineal de P elementos de bases de excitación w_p^{F0} y E bases elementales de filtros w_p^Φ como se muestra a continuación:

$$s_t = s_t^\Phi \bullet s_t^{F0} = \sum_{e=1}^E h_{et}^\Phi w_e^\Phi \sum_{p=1}^P h_{pt}^{F0} w_p^{F0} \quad (9)$$

donde h_{et}^Φ y h_{pt}^{F0} son las ganancias no negativas y $\bullet$ denota el producto de Hadamard. Las bases de excitación w_p^{F0} representan la colección discreta de sonidos con la que cada señal

se construye y cada una es modulada por la combinación de bases w_e^ϕ . Ya que el sistema está hecho para representar vocales, cada uno de los vectores w_p^{F0} representa la señal glotal correspondiente a una frecuencia fundamental o tono. Las bases w_p^{F0} se generan usando un modelo de fuente glotal para generar un diccionario de excitación de tonos. Las bases del filtro w_e^ϕ tienen que ser capaces de representar la forma de la señal. Por tanto, las bases de excitación w_p^{F0} se pueden agrupar en una matriz como $W^{F0} = [w_1^{F0}, \dots, w_p^{F0}]$, y las bases w_e^ϕ en una matriz tal que $W^\phi = [w_1^\phi, \dots, w_E^\phi]$. El modelo se puede representar como una matriz de la forma:

$$S = (W^\phi H^\phi)(W^{F0} H^{F0}) \quad (10)$$

Donde las matrices de ganancia $[H^\phi]_{et} = h_{et}^\phi$ y $[H^{F0}]_{pt} = h_{pt}^{F0}$ nos da la descomposición de la fuente de habla dentro de los diccionarios dentro de W^ϕ y W^{F0} , y las amplitudes son estimadas desde la señal de entrada mientras que los diccionarios se mantienen fijos.

Para la caracterización de las bases para el ruido, se verán en la parte de implementación de software, ya que se ha utilizado el mismo método para su caracterización y detección.

2.1.2.2 *Análisis de datos de microarrays [2]*

Se hace uso de un algoritmo de NMF que puede controlar el grado de diseminación en el coeficiente de la matriz H a través de mínimos cuadrados alternando restricciones de no negatividad para analizar datos de microarrays. Se aplica NMF para el entrenamiento de los datos en una matriz donde las filas de los microarrays representan, en este caso en concreto, genes y las columnas representan los experimentos, partiendo de la ecuación (1).

Al introducir datos de estas enfermedades, se puede llevar a cabo una identificación de una enfermedad en cuestión con buena probabilidad de acierto siendo posible extrapolarlo a otros tipos de cáncer o incluso análisis de datos de expresión génica.

Se pretende a partir de una cantidad importante de información, buscar patrones que tengan probabilidad de repetirse. Estos patrones serán nuestras bases de caracterización y nos permitirán al analizar otro tipo de datos teniendo como resultado excitación de las bases obtenidas anteriormente, permitiéndonos conocer si existe o no presencia de las enfermedades caracterizadas en los nuevos datos analizados.

2.1.2.3 Separación de fuentes habladas grabadas en un sólo canal [3]

La separación de fuentes de voz grabadas con un solo micrófono trae consigo problemas por no tener información espacial de las situaciones en las que se encuentran nuestras fuentes. El método consiste en un sistema no supervisado que aprenda la representación escasa de los datos. Se utiliza para el aprendizaje de los diccionarios del cuerpo del habla para separar el flujo de audio en sus componentes.

Pariendo de la ecuación (1) pero modificada su nomenclatura, $Y = DH$, donde el diccionario es D , y la matriz de activación H . Se asume que la mezcla de la señal es la suma de las R fuentes de señal:

$$Y = \sum_i^R Y_i \quad (11)$$

Las bases de la mezcla de la señal es la concatenación de los diccionarios fuente, $D = [D_1 \dots D_i \dots D_R]$, y el código completo de la matriz es la concatenación de los diferentes códigos de fuentes-individuales, $H = [H_1^T \dots H_i^T \dots H_R^T]^T$. Mediante el forzado de la escasez de la matriz H es posible separar Y in sus diferentes fuentes si los diccionarios son los suficientemente diversos.

NMF básico no permite definir una buena solución en el caso de definir diccionarios completos cuando las restricciones de no negatividad son insuficientes para tener una solución concreta. En este caso, se utiliza SNMF (sparse non-negative matrix factorization) la cual optimiza el coste de la función:

$$E = \|Y - DH\|_F^2 + \lambda \sum_{ij} H_{ij} \quad D, H \geq 0 \quad (12)$$

Donde D es la columna normalizada de la matriz diccionario. Para aprender los diccionarios de manera individual para cada hablante se aplica SNMF y para separar las mezclas de voz se mantienen los diccionarios fijos y se van actualizando sólo la matriz de activación H .

Finalmente, la voz se separará mediante reconstrucción de las partes de las descomposiciones usadas por los diccionarios.

2.1.2.4 Transcripción de música polifónica [4]

Se parte de pasajes musicales polifónicos compuestos por notas que muestran un perfil espectral armoniosamente fijo. Se puede modelar el contenido de audio por una transformación base lineal y un uso de NMF para estimar el perfil espectral y la información temporal de cada nota. El sistema no está basado en conocimiento, sino que aprende las posibles notas por observación.

Partiendo de la magnitud espectral de un segmento del sonido

$$s(t) = g(\alpha t)\sin(\gamma t) + g(\beta t)\sin(\delta t) \quad (13)$$

Donde $g()$ es una función con un periodo 2π y $\alpha, \beta, \delta, \gamma$ son escalares arbitrarios con α y β significativamente pequeños frente a δ y γ . Tras analizar el tiempo de duración dependiente del espectro de magnitud $x(t) = \|DFT([s(t) \dots s(t+L)])\|$. Todos los valores de x se almacenan como columnas en una matriz no negativa $X_{M \times N}$, donde N denota el número total de espectros y M es el número de sus frecuencias.

Establecida la matriz anterior, se aplica NMF para detectar en qué frecuencias se tiene más energía donde el espectrograma resulta ser altamente redundante y gracias a ella se puede realizar una descripción más compacta para optimizar la función de coste.

La posibilidad de detectar en el espectro donde se centra la frecuencia, permite realizar un aprendizaje de las notas presentes ya que es un escenario redundante y los armónicos no contienen toda la potencia como las frecuencias fundamentales de ellas.

2.1.2.5 Fusión de imágenes pancromáticas y multiespectrales capturadas mediante satélite [6]

En esta ocasión se trata de mejorar el sistema de fusión de imágenes capturadas vía satélite para evitar la distorsión espectral y la necesidad de introducir una descomposición de frecuencias compleja y como resultado, teniendo un alto coste computacional y mucho tiempo de procesado.

Aplicando NMF junto con otra técnica denominada Hyperspherical Color Sharpening (HCS), donde el algoritmo utilizado es capaz de mantener las características espectrales de las imágenes en el proceso de reconstrucción, evitando una descomposición de frecuencias, y el nivel de detalle espacial se incorpora perfectamente, consiguiendo un sistema de fusión de imágenes rápido y preciso en tiempo real.

Partiendo de la ecuación (1) y aplicando las reglas definidas en las ecuaciones (4) y (5) se obtiene una descomposición de los datos no negativos y los vectores base, obtenidos por medio de un algoritmo de iteración multiplicativa de aprendizaje, contienen las características del objeto estudiado.

Utilizando NMF, se extrae de imágenes multiespectrales una componente “I”, la cual, tendrá una respuesta espectral consistente con la imagen pancromática tanto como sea posible. Por tanto, la componente “I” extraída mediante NMF, la imagen pancromática “P” y las imágenes multiespectrales de 8 bandas “M” constituyen la matriz V que será descompuesta. El método de extracción de la componente “I” por medio de NMF sería el siguiente:

$$[P, M_1, M_2, \dots, M_8] = WH \quad (14)$$

Donde $P, M_1, M_2, \dots, M_8$ son las imágenes multiespectrales pancromáticas y de 8 bandas. La componente “I” extraída debe coincidir con la imagen pancromática y la matriz W es un vector columna donde la imagen restaurada desde W es la componente obtenida “I”.

2.1.3 Detección de insectos mediante captación de señales

En este apartado, se van a explicar algunas formas de detección e identificación de insectos que actualmente se están utilizando.

2.1.3.1 Captura y procesamiento de sonidos [18]

Se realiza la captura de los sonidos producidos por diversos insectos en campo y se transforma en audio digital para llevar a cabo un análisis de la señal. Para ello, se realiza al audio digitalizado la transformada rápida de Fourier. De esta manera tendremos una representación de la respuesta en frecuencia del audio capturado y para poder tener unas frecuencias más fiables, se aplica un filtro paso banda. Una vez obtenidas las frecuencias más representativas se crea una red neuronal para analizar los patrones específicos de cada especie y llevar a cabo la identificación.

2.1.3.2 Registro de vibraciones producidas por insectos (I) [13]

Para la identificación de insectos muy pequeños, los cuales no producen un sonido captable por el ser humano, se utilizan transductores magneto-inductivos.

El proceso es simple: se coloca un imán en el tallo de una planta donde haya insectos. A continuación se coloca una bobina situada muy cerca del imán sin que esté en contacto con él. De esta forma, al producirse pequeñas vibraciones debidos a los insectos recorriendo la planta, se traspasarán a la bobina y se registrará a través de un sistema capaz de digitalizar la señal.

La complejidad del sistema radica en la colocación del imán ya que una colocación errónea podría causar registros incorrectos.

2.1.3.3 Registro de vibraciones producidas por insectos (II) [13]

Mediante el uso de un instrumento denominado Vibrómetro Láser, se mide la velocidad producida en una superficie gracias a los cambios de frecuencia en el láser reflejado. El uso de esta técnica se basa en la teoría del efecto Doppler y sus resultados son muy fiables, ya que al no estar en contacto con la superficie estudiada no se producen posibles contaminaciones de los resultados. Este sistema es capaz de medir zonas específicas como partes de los insectos para llegar a identificarlos, como por ejemplo las antenas de los Ortópteros.

2.2 Base de datos de insectos utilizada

Se ha utilizado una base de datos de audios de diferentes insectos a partir de los cuales, se llevará a cabo un procesamiento de la señal para la caracterización de cada tipo de insecto, como veremos más adelante. El sonido se ha extraído de una base de datos alojada en internet, donde las grabaciones fueron grabadas mediante un micrófono de condensador Avisoft-UltraSound Gate 116. Es un micrófono dedicado a grabar ultrasonidos en el aire.



Figura 1 Micrófono Avisoft-UltraSound Gate 116

Las grabaciones, en algunos casos han sido modificadas ralentizando la velocidad de reproducción por un factor específico para hacerlas audibles. Todas las grabaciones de la base de datos han sido realizadas en campo [17].

En la base de datos utilizada, se encuentran registrados sonidos de insectos pertenecientes a diversos órdenes, tales como Acrididae (saltamontes) o Gryllidae (grillo) pertenecientes a la orden de los Ortópteros; Sphingidae (mariposa) de la orden de los Lepidópteros y Cicadae (cigarra) de la orden de los Hemípteros.

Para obtener los sonidos de los insectos, es necesario que éstos lo produzcan, por lo cual se va a proceder a comentar el aparato generador de sonido característico de algunos de los insectos incluidos en nuestra base de datos.

En la siguiente tabla se muestran los insectos utilizados para realizar todas las pruebas. La segunda columna corresponde con el nombre del fichero de audio utilizado y la tercera, el nombre científico del insecto en cuestión.

Nº insecto	Nombre del fichero	Nombre del insecto
1	Acherontia_atropos.wav	Acherontia atropos
2	Acheta_domesticus.wav	Acheta domesticus
3	bangi_i1.wav	Heuschrecke
4	chort_brun.wav	Chorthippus brunneus
5	cicada_barbara.wav	Cicada barbara
6	cicada_cretensis.wav	Cicada cretensis
7	cicada_orni.wav	Cicada orni
8	con_dor.wav	Conocephalus dorsalis
9	g1.wav	Gryllus
10	gryllotalpa_gryllotalpa.wav	Gryllotalpa gryllotalpa
11	gryllus.wav	Gryllus
12	Gryllus_bimaculatus.wav	Gryllus bimaculatus
13	h1.wav	Cicadatra atra
14	Heimchen.wav	Acheta domesticus
15	Leptophyes_punctatissima.wav	Leptophyes punctatissima

16	metr_roe.wav	Metrioptera roeselii
17	Metrioptera roeselii.wav	Metrioptera roeselii
18	Stethophyma_grossum.wav	Stethophyma grossum
19	u1.wav	Leptophyes punctatissima
20	u2.wav	Platycleis albopunctata
21	u3.wav	Decticus verrucivorus
22	u4.wav	Myrmeleotettix maculatus
23	u5.wav	Chorthippus mollis
24	u6.wav	Metrioptera bicolor
25	u7.wav	Chorthippus vagans

Tabla 1 Base de datos de insectos

2.2.1 Cicadae (Cigarra)

La cigarra común no produce su sonido característico en todas las fases de su vida.

En este caso, su ciclo puede variar entre 13 y 17 años para pasar por todos sus estados, desde larva hasta ser adulto. En éste último tramo de su vida, es donde produce el sonido característico que casi todos nosotros conocemos.

Las cigarras [14] podemos escucharlas en verano cuando nos encontramos en días especialmente calurosos. Dicho sonido, es producido sólo por los machos, ya que es la época de apareamiento y pretenden llamar a las hembras para reproducirse.

Su sistema de producción de sonido, sin entrar en términos muy técnicos, es un órgano situado a cada lado del abdomen llamado “estrídulador”. Está formado por varias membranas llamadas timbales y cavidades, o sacos, de aire que harán la función de caja de resonancia. Para producir el sonido, cuyo nivel es capaz de llegar a los 158 dB, utilizan la retracción de un músculo específico para ello.

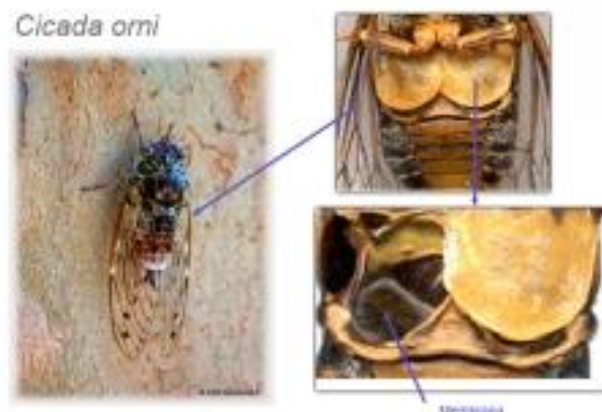


Figura 2 Órgano estrídulador Cicadae

2.2.2 *Acrididae* (Saltamontes)

La forma de producir sonido de los saltamontes [15] es totalmente diferente a las comentadas anteriormente. De nuevo, el canto de estos insectos pertenece a los machos para atraer a las hembras y poder llegar a reproducirse.

En este caso, el órgano de producción del sonido es del tipo fémoro-alar donde al frotar el fémur posterior (compuesto de pequeñas púas) contra una vena situada en sus alas, se produce el sonido común en ellos.

Destacar el comportamiento de los saltamontes al variar el sonido que producen dependiendo del ruido que los rodea para, así, poder destacar y conseguir atraer a las hembras.

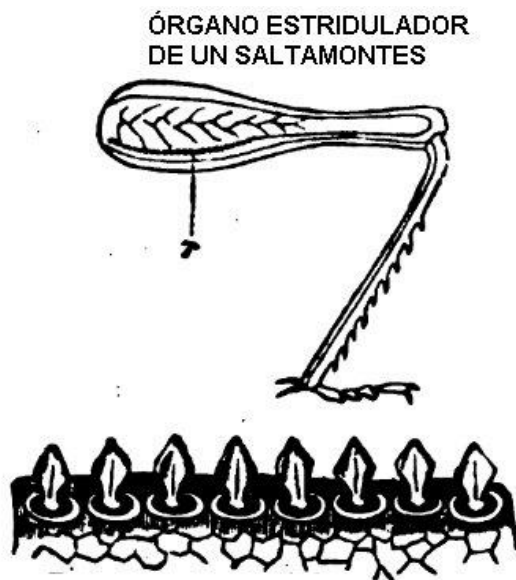


Figura 3 Órgano estridulador *Acrididae*(I)

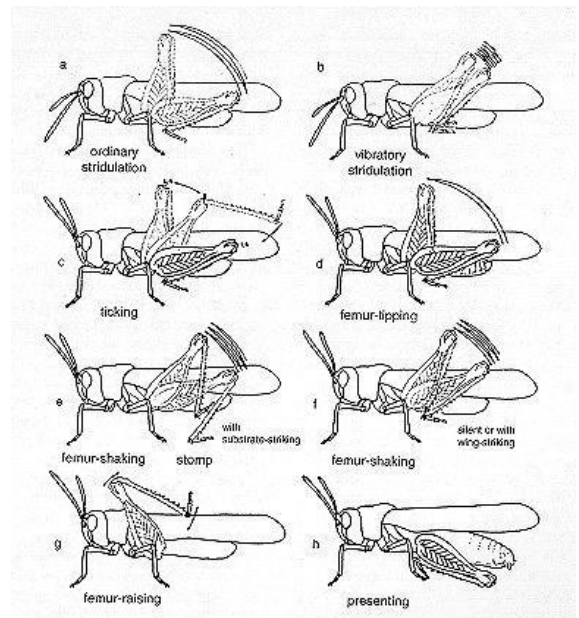


Figura 4 Órgano estridulador *Acrididae* (II)

2.2.3 *Gryllidae* (Grillo)

Los grillos producen su sonido de manera muy similar a la de los saltamontes ya que se realiza por frotamiento. Sin embargo, éstos utilizan sus alas en vez de sus extremidades traseras como en el caso de los saltamontes. De nuevo, el sonido producido tiene como fin cortejar a las hembras para poder reproducirse y son capaces de modular los tonos emitidos dependiendo de los factores externos como la temperatura o incluso para llamar la atención, por encima de otro espécimen, de las hembras.



Figura 5 Gryllidae

2.3 Base de datos de ruido utilizada

La base de datos de ruido con la que se han realizado las pruebas ha sido “NOIZEUS”. Es una base de ruido desarrollado para la comparación de algoritmos de mejora del habla para investigación. Está compuesta por 30 frases ruidosas producidas por 3 hablantes masculinos y 3 hablantes femeninos corrompidos por ocho tipos de ruido del mundo real con diferentes relaciones señal/ruido (SNR). El ruido introducido en estos ficheros ha sido ruido de tren, coche, sala de exposiciones, restaurante, calle, aeropuerto y ruido de estación de tren. Para cada tipo de ruido, tenemos 30 ficheros diferentes con diferentes SNR. En total se tienen 960 ficheros de ruido.

Los ficheros son modificados para simular las características de frecuencia en la recepción de manera telefónica, por lo que se ha filtrado la señal mediante un filtro IRS (Intermediate Reference System).

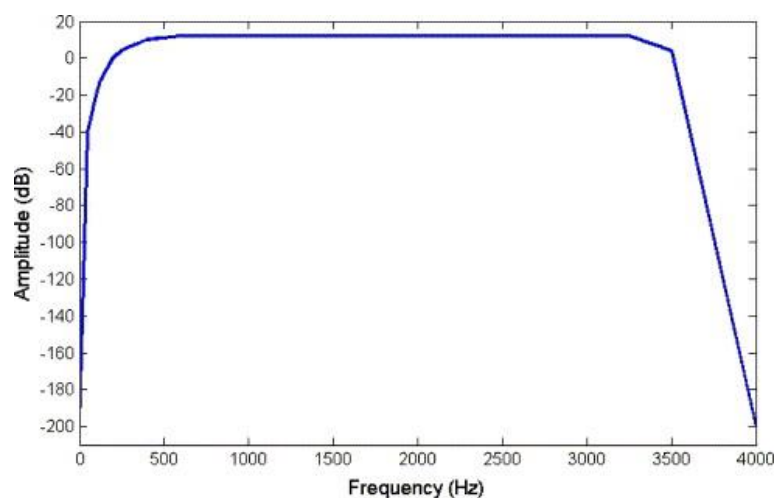


Figura 6 Respuesta en frecuencia filtro IRS

Para nuestras pruebas, se debía modificar la base de datos de ruido ya que contenían información del habla y se buscaba el ruido aislado para combinarlo con nuestra base de datos de insectos, por ello, se ha eliminado la señal del habla de los ficheros. Ya que se tenían los ficheros del habla, se ha procedido a la resta de ambas señales (ruido con habla y habla aislada) para así obtener las señales de ruido buscadas para nuestras pruebas.

En un siguiente paso, se ha procedido a unir todos los ficheros de ruido del mismo tipo con misma SNR. Para no introducir alta frecuencia en la unión de los ficheros, se ha procedido a multiplicar las muestras iniciales y finales de los distintos ficheros con una señal rampa. Las muestras iniciales del fichero se multiplican con una señal rampa positiva que parte desde 0 hasta 1 del tamaño de muestras deseado.

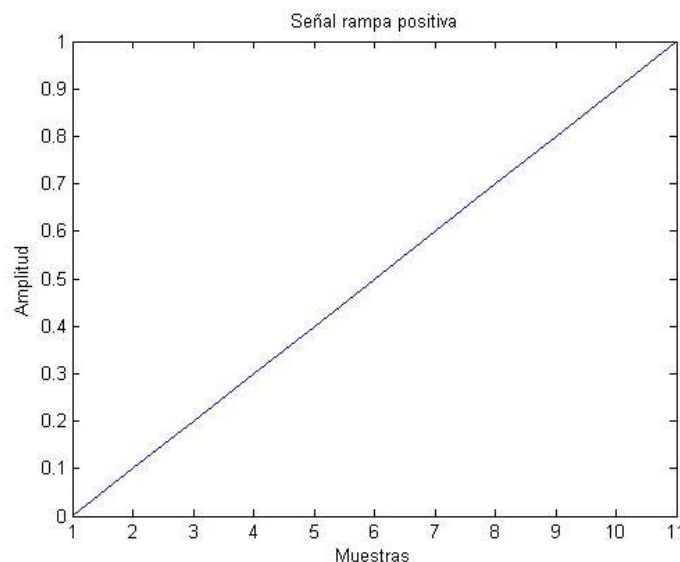


Figura 7 Señal rampa positiva de 10 muestras

Tras realizar la combinación de la señal rampa positiva (en el caso superior de 10 muestras) con las primeras muestras del fichero de ruido, conseguimos tener todos los ficheros de ruido partiendo desde el valor 0 facilitando la unión entre todos ellos.

Para las muestras finales de los ruidos, se realiza el mismo proceso con la diferencia de utilizar una señal rampa negativa, de manera que las muestras finales del audio (última muestra) tiendan a 0 y coincida con la primera muestra del siguiente fichero de audio.

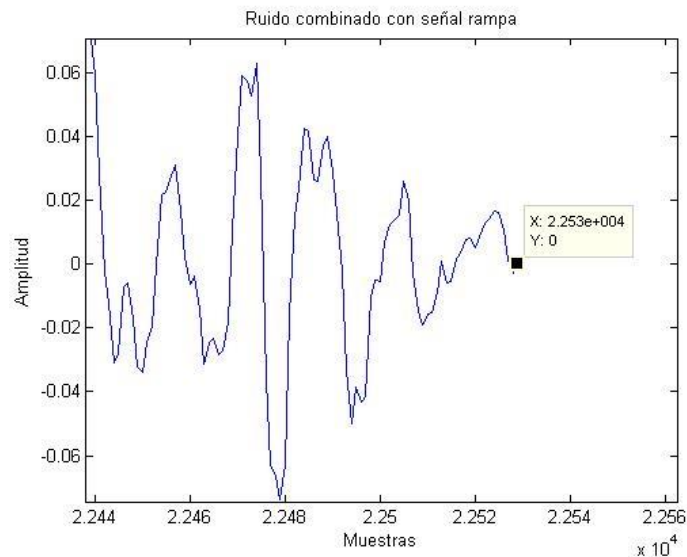


Figura 8 Muestras finales del fichero 1 de ruido Airport multiplicado por señal rampa negativa

En la gráfica superior se observa como la última muestra del fichero toma el valor 0. En la siguiente gráfica se muestra el inicio de la señal con la primera muestra en 0 de la misma manera.

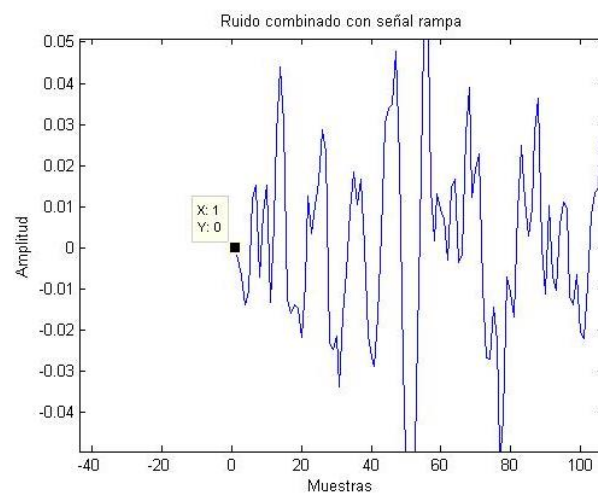


Figura 9 Muestras iniciales del fichero 1 de ruido Airport multiplicado por señal rampa positiva

En esta última gráfica, se muestra la unión de dos ficheros de ruido concatenados para generar un solo fichero de ruido. Como vemos, la unión se produce con valor 0 para evitar introducir interferencias de frecuencia dentro del fichero generado.

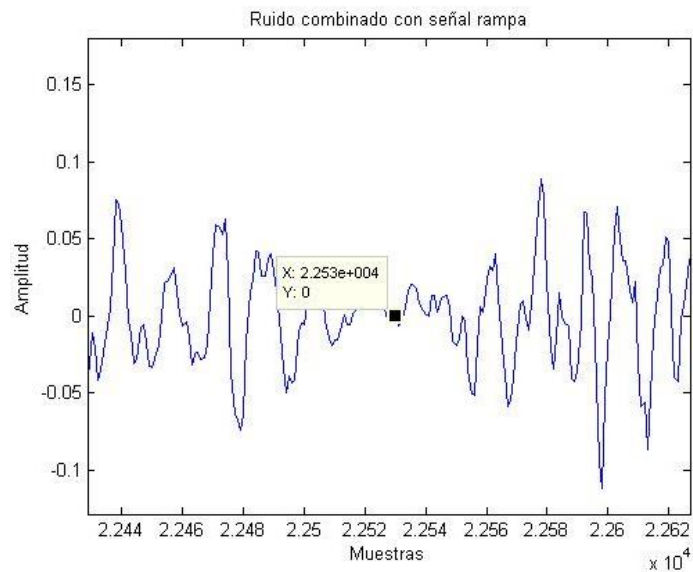


Figura 10 Unión ficheros 1 y 2 de ruido Airport

2.4 Modelo utilizado

Se ha utilizado un software modificado basado la mejora de detección de voz [1]. El software implementa la factorización de matrices no negativas (NMF) para la extracción de fuentes y comprobar las contribuciones de cada una de ellas, a partir de un fichero de audio.

Se ha procedido a la adaptación del software original para un correcto funcionamiento en la identificación de insectos. En concreto, el software inicial tenía en cuenta la frecuencia de la voz, producido por la glotis donde se encuentran las cuerdas vocales. En nuestro caso, se ha eliminado esta dependencia ya que los insectos no cuentan con dicha función y por ende los resultados obtenidos al usar el software original serían erróneos.

A continuación, veremos el desglose del modelo utilizado y como se caracteriza matemáticamente cada una de las partes partícipes en este

2.4.1 Principios básicos de la separación de fuentes con NMF [1]

2.4.1.1 Uso de NMF para la separación de fuentes

Los algoritmos de separación que hacen uso de NMF se basan en la composición de modelos de señal donde los espectrogramas de cada fuente están modelados como combinaciones variantes en el tiempo de funciones base específicas, que denominaremos bases, y donde todos los elementos que componen el modelo son no negativos. s_t es el vector espectral de la fuente (insecto) en la trama t . En su versión simplificada, el modelo s_t es una combinación lineal de las funciones base w_j^s como:

$$s_t = \sum_{j=1}^J h_{jt}^s w_j^s \quad (15)$$

donde J es el número de bases de sonido de insectos y h_{jt}^s es la ganancia de las j -ésima función base de la trama t . Del mismo modo, cada vector espectral de ruido b_t puede expresarse como una combinación lineal de K funciones bases de ruido w_k^b con ganancia h_{kt}^b . El espectrograma de entrada se puede representar como una combinación del sonido del insecto y las bases de ruido:

$$X = W^S H^S + W^B H^B = WH \quad (16)$$

donde $W^S = [w_1^s, \dots, w_J^s]$, $W^B = [w_1^b, \dots, w_K^b]$, $[H^S]_{jt} = h_{jt}^s$ y $[H^B]_{kt} = h_{kt}^b$. Las matrices W^S y W^B serán nuestros diccionarios de fuentes que contienen los patrones a partir de los que se puede reconstruir cada trama. En cambio, las matrices H^S y H^B determinan cómo se combinan en cada trama la fuentes del audio. El conjunto de bases W y H nos permiten reconstruir la señal completa.

El error de reconstrucción se obtiene de la función $D(X, WH) = \sum_{ft} d([X]_{ft}, [WH]_{ft})$, donde $d(a, b)$ ofrece la divergencia entre los dos escalares. El cálculo de coste de divergencia se puede calcular mediante la siguiente expresión de β -divergencia:

$$d_{\beta}(a, b) = \begin{cases} \frac{1}{\beta(\beta-1)}(a^{\beta} + (\beta-1)b^{\beta} - \beta ab^{\beta-1}), & \beta \in \mathbb{R}^+ \setminus \{0,1\} \\ a \log \frac{a}{b} - a - b, & \beta = 1 \\ \frac{a}{b} \log \frac{a}{b} - 1, & \beta = 0 \end{cases} \quad (17)$$

donde los valores típicos usados en NMF son $\beta = 2$, $\beta = 1$ y $\beta = 0$.

2.4.2 Descomposición del sonido del insecto y separación del ruido [1]

Para realizar el análisis de la señal de entrada se divide, en un principio, en tramas de 64 ms de longitud con un solapamiento del 50% y se aplica un enventanado con la ventana de Hamming.

Se han realizado diversas pruebas para comprobar cuál es el mejor valor de solapamiento y tiempo de trama y se ha llegado a una conclusión de que los valores utilizados son los más óptimos debido a un mayor porcentaje de éxito en el reconocimiento de los insectos.

A continuación, se calcula la STFT (Short-time Fourier transform) y se discretiza con una resolución de $\frac{1}{4}$ de semitono mediante la integración de los valores de magnitud dentro de cada banda con $\frac{1}{4}$ de semitono de ancho.

$$x_{ft}^0 = \sum_{F(f)}^{F(f+1)} |X(\Omega, t)| d\Omega \quad (18)$$

Donde $X(\Omega, t)$ es la STFT en la trama t , Ω es la frecuencia en Hz y $F(f) = 32.7 * 2^{(f-1)/48}$ con $F = 380$ puntos de frecuencia.

El espectrograma se normaliza teniendo en cuenta β , usado en la β -divergencia, y el tamaño del espectrograma se calcula de la siguiente manera:

$$x_{ft} = \frac{x_{ft}^0}{\left(\frac{1}{FT} \sum_{f=1}^F \sum_{t=1}^T x_{ft}^0 \right)^{1/\beta}} \quad (19)$$

Donde F es el número de elementos de frecuencia y T el número de tramas.

2.4.2.1 Modelo de señal

El espectrograma de entrada se modela por la suma de 4 contribuciones:

$$X \approx \hat{X} = W^\Phi Z^\Phi H^\Phi + W^N Z^N H^N + W^I Z^I H^I + W^H Z^H H^H \quad (20)$$

El primer término representa el sonido del insecto. El resto de términos representan el ruido como suma de 3 tipos diferentes. Cada uno está expresado por una matriz diferente factorizada en un conjunto de funciones base y sus activaciones.

Cada tipo de ruido está caracterizado por un comportamiento espectro-temporal y sus bases son caracterizadas individualmente. Los diferentes tipos de ruido son definidos de la siguiente manera:

1. El primer tipo de ruido (ruido suave), está caracterizado por un espectrograma con variaciones pequeñas entre valores próximos. Es modelado por la combinación de bases con una forma espectral lisa en frecuencia y la activación en el tiempo se produce por variaciones suaves. Los términos que lo definen son $W^N Z^N H^N$.
2. El segundo tipo hace referencia al ruido impulsivo. Se produce un pico de energía en un intervalo corto de tiempo y se puede modelar como suavidad espectral en la frecuencia y activaciones en instantes de tiempo. Sus términos son $W^I Z^I H^I$.
3. El último término se define como el ruido armónico, donde tenemos interferencias de banda estrecha y se activan en tramas consecutivas. Se puede modelar por una escasez en la frecuencia y su ganancia tiene pequeñas variaciones en el tiempo. Se denota por los términos $W^H Z^H H^H$.

2.4.3 Descomposición del algoritmo [1]

El set de parámetros del modelo $\theta = [H^\Phi, W^N, H^N, W^H, H^H]$ se estima minimizando el coste de la función

$$\mathcal{C}(\theta) = D(X, \hat{X}) + C_N(W^N, H^N) + C_I(W^I, H^I) + C_H(W^H, H^H) \quad (21)$$

Donde el error de reconstrucción de $D(X, \hat{X})$ está medido por la β -divergencia. A continuación, se explicarán las restricciones de cada tipo de ruido.

2.4.3.1 Restricciones de ruido suave

Los parámetros W^N, H^N son los encargados de capturar el tipo de ruido con un espectrograma suave como pueden ser los ruidos rosas, blancos y de conversaciones. La función $SSM()$ denota la suavidad espectral de la matriz W^N , expresado como sumas cuadráticas de diferentes frecuencias consecutivas.

$$SSM(W^N) = \sum_{r=1}^{R_R} \sum_{f=2}^F (w_{fr}^N - w_{f-1,r}^N)^2 \quad (22)$$

donde $[W^N]_{fr} = w_{fr}^N$. De manera similar la suavidad temporal de la matriz H^N es medida mediante la función $TSM()$.

$$TSM(H^N) = \sum_{r=1}^{R_R} \sum_{t=2}^T (h_{tr}^N - h_{r,t-1}^N)^2 \quad (23)$$

Donde $[H^N]_{rt} = h_{rt}^N$. Según las restricciones anteriores el coste de las bases y ganancias del ruido suave viene dado por:

$$C_N(W^N, H^N) = \alpha \gamma_W^N SSM(W^N) + \alpha \gamma_H^N TSM(H^N) \quad (24)$$

Donde α es un coeficiente positivo que pesa la importancia de las restricciones de suavidad y del costo global, y γ_W^N y γ_H^N son la constante de normalización.

2.4.3.2 Restricciones de ruido impulsivo

El ruido impulsivo aparece en pequeños intervalos del tiempo donde se concentra toda la energía y es posible que se repitan. Como ejemplos tenemos el ruido de máquinas o el de un disparo por ello, necesita una caracterización de corta duración en el tiempo ya que se activarán en instantes aislados del tiempo. Los parámetros que lo caracterizan son W^I, H^I y el coste de la función viene dada por:

$$SP(H^I) = \sum_{r=1}^{R_R} \sum_{t=1}^T h_{rt}^I \quad (25)$$

Donde $[H^I]_{rt} = h_{rt}^I$. La suavidad de los vectores base en W^I son controladas por la función $SSM()$ y el término que contiene las bases y ganancia del ruido impulsivo se expresa de la siguiente manera:

$$C_I(W^I, H^I) = \alpha \gamma_W^I SSM(W^I) + \lambda \gamma_H^I SP(H^I) \quad (26)$$

Donde λ es un peso positivo para equilibrar la importancia de las medidas de escasez en el coste global de la función, α determina el peso de los términos de suavidad y γ_W^I y γ_H^I son las constantes normalizadas.

2.4.3.3 Restricciones de ruido de armónico

Este tipo de ruido viene caracterizado por W^H y H^H . Las bases espectrales serán muy escasas con algunos elementos activos mientras que sus ganancias serán relativamente suaves en el tiempo.

Se aplicará la función de coste $SP()$ a los parámetros W^H y H^H y se aplicará la función $TSM()$ a las ganancias H^H para tener estabilidad temporal. El coste total para el ruido armónico será:

$$C_H(W^H, H^H) = \alpha \gamma_W^H SP(W^H) + \lambda \gamma_H^H TSM(H^H) \quad (27)$$

Donde α y λ son los parámetros específicos de suavidad y escasez.

2.4.3.4 Estimación de parámetros

El algoritmo de estimación está basado en la estrategia de gradiente multiplicativa donde las matrices desconocidas del modelo son inicializadas en un primer momento de manera aleatoria con números positivos y se actualizan alternativamente en cada iteración con reglas de actualización multiplicativas.

La matriz de ganancia H^Φ describe la descomposición del sonido del insecto restringido sólo por el error de reconstrucción. Para reducir la β -divergencia entre X y el modelo $\hat{X}$, se aplica la siguiente regla de actualización:

$$H^\Phi \leftarrow H^\Phi \frac{(W^\Phi Z^\Phi)^T (\hat{X}^{\beta-2} \bullet X)}{(W^\Phi Z^\Phi)^T \hat{X}^{\beta-1}} \quad (28)$$

donde $\bullet$ es el producto de Hadamard, $(\cdot)^T$ es el operador traspuesta y todas las fracciones y exponenciales son aplicadas elemento a elemento.

Las matrices de ruidos están restringidas adicionalmente por los términos de regularización. Las actualizaciones de las reglas para estas matrices pueden derivarse sabiendo la regla de actualización multiplicativa para una matriz genérica A como

$$A \leftarrow A \bullet \frac{\nabla_A^- C(\theta)}{\nabla_A^+ C(\theta)} \quad (29)$$

donde $\nabla_A C(\theta)$ es la derivada parcial de la función objetivo $C(\theta)$ respecto de A .

2.4.3.5 Síntesis de la señal

Una vez estimados los parámetros, cada componente es sintetizada en el dominio del tiempo. Si $\hat{S} = (W^\Phi Z^\Phi H^\Phi)$ es el espectrograma estimado del sonido del insecto, se aplica un filtro Wiener, siendo su máscara Wiener espectro-temporal M^S y denota de la siguiente manera:

$$M^S = \frac{\hat{S}^2}{\hat{S}^2 + (W^N Z^N H^N)^2 + (W^I Z^I H^I)^2 + (W^H Z^H H^H)^2} \quad (30)$$

RESULTADOS Y DISCUSIÓN

El desarrollo del Trabajo de Fin de Grado se ha dividido en diferentes fases. Inicialmente se ha tenido que realizar un estudio y familiarización con el software facilitado además de la obtención de las bases de datos a utilizar y posterior adaptación del software para poder desarrollar el objetivo planteado.

3.1 Adaptación del software a las necesidades del TFG

Tras conseguir el funcionamiento correcto de nuestro sistema con ficheros de sonidos de insectos, se ha llevado a cabo el entrenamiento de las bases que caracterizan a cada insecto a partir de una base de datos con distintos ficheros de sonidos de éstos. Se consigue de esta manera generar un diccionario donde se encuentran todas las bases obtenidas y tendremos todos los insectos caracterizados.

Una vez generado el diccionario de bases de insectos, se ha procedido a la comprobación de capacidad de reconocimiento, en diferentes niveles, que es capaz de ofrecer el software. En primera instancia, se ha hecho una primera prueba introduciendo los propios ficheros de insectos originales, a partir de los que se han generado las bases, para comprobar que efectivamente se identifica al insecto en cuestión. Una vez completada, se ha procedido a la adición de diferentes tipos de ruido. En un primer caso, se ha añadido ruido blanco gaussiano con diferentes relaciones señal a ruido (desde ahora SNR) para comprobar la robustez que se logra alcanzar. Decir que este tipo de ruido, se puede caracterizar simplemente con unos de los términos de ruido implementado en el software, en concreto el término de ruido pseudoestacionario, y no es una situación real, debido a que rara vez se dan los sonidos a identificar con únicamente este tipo de ruido.

En la etapa final de esta fase, se ha desarrollado un código capaz de añadir diferentes tipos de ruido de situaciones más reales. Se ha procurado adaptar las señales para evitar introducir tonos de alta frecuencia y modificar las posibles lecturas e inducir a error. En concreto, se ha utilizado una base de datos denominada “Noizeus”, la cual, además del ruido a utilizar, contenía información no necesaria (como ya se ha explicado en el apartado Base de datos de ruido). Preparada las bases de datos de insectos y ruido, estamos en condiciones de proceder a la combinación de ambos para realizar las pruebas pertinentes y comprobar el funcionamiento del sistema NMF aplicado a la identificación de insectos.

3.2 Actualización del software

El software desarrollado, ha sido muy modular, con diferentes funciones para realizar los procesos correspondientes y poder depurar en cualquier momento diferentes parámetros para el correcto funcionamiento o posible mejora.

Para poder proceder a ejecutar el software para la obtención de resultados, se ha tenido que actualizar para ser compatible con el entorno MATLAB R2016a, ya que ha sido esta versión, la responsable de ejecutar el programa para obtener los resultados. La actualización del software, para hacerlo compatible con una versión superior de MATLAB, ha sido necesaria debido a la eliminación y corrección de diferentes funciones incluidas en la toolbox y modificación del mecanismo de ejecución en los algoritmos en versiones posteriores a MATLAB R2013a. Por ello, para la elaboración del código se ha utilizado el entorno MATLAB R2011b y para la ejecución del mismo, se ha actualizado para la versión MATLAB R2013a o superior.

3.3 Resultados obtenidos

Una vez completado el código y comprobado su correcto funcionamiento, probando el sistema de reconocimiento con los ficheros originales de los diferentes insectos, se han realizado diversas pruebas con diferentes tipos de ruido para obtener y analizar unos resultados completos a la hora de sacar unas conclusiones.

Como parámetros de entrada en nuestro sistema, nos encontraremos con el insecto a identificar, el tipo de ruido añadido, el nivel de SNR, el número de bases para caracterizar el insecto y los 3 tipos de ruido implementado en el software y con dos últimos valores responsables del resultado de excitación del número de bases que caracterizan a cada insecto, parámetro de smoothness (suavidad) y sparseness (escasez).

Para decidir qué insecto es el que está produciendo el ruido, se realiza la suma de cada una de las filas que compone la matriz de activación (H_f) y se calcula, teniendo en cuenta en número de bases utilizadas, cuál es la combinación con mayor valor (mayor peso en la activación) y se determina como el insecto resultante detectado.

3.3.1 Resultados obtenidos añadiendo ruido blanco a los ficheros de sonidos de insectos

A la base de datos de los insectos, se le ha añadido ruido blanco gaussiano para comprobar cuan robusto es el sistema. Los parámetros utilizados en este proceso han sido 5 bases para caracterizar tanto los insectos como los tipos de ruido y como valores de

sparseness y smoothness se ha utilizado un valor igual a 1 (como veremos más adelante es una de las combinaciones óptimas), ya que como veremos más adelante es la combinación con la que se tienen mejores resultados. Se han comprobado las posibles SNR mínimas donde es capaz de identificar correctamente el sistema los diferentes insectos.

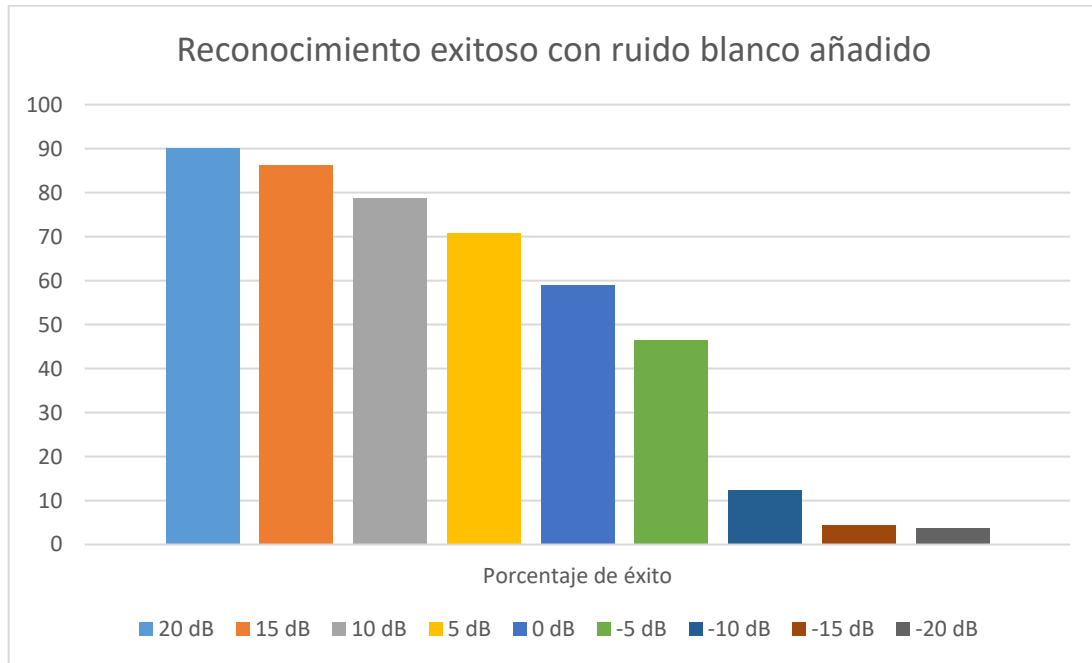


Figura 11 Porcentaje de éxito para diferentes niveles de SNR de todos los insectos añadiendo ruido blanco gaussiano. El total de operaciones por cada nivel ha sido de 161. La configuración usada ha sido 5 bases de caracterización para el insecto y los diferentes tipos de ruido con un valor para sparseness y smoothness igual a 1.

Los resultados nos muestran un buen funcionamiento por encima de 0 dB de SNR. Por debajo de este nivel los errores se hacen más notables, sin embargo de forma general, se es capaz de llegar a una SNR de -5 dB y con determinados insectos, incluso se puede alcanzar una SNR de -10 dB.

Estos niveles de SNR son muy esperanzadores, sin embargo, realizar pruebas con únicamente ruido blanco gaussiano no es lo más ideal ya que si se quiere aplicar el sistema al mundo real, será muy difícil capturar los sonidos de los insectos con únicamente este tipo de ruido.

Por ejemplo, en la figura 11 tenemos un insecto (insecto 1) al que se le ha añadido ruido blanco con un nivel de SNR de -15 dB y se detecta correctamente.

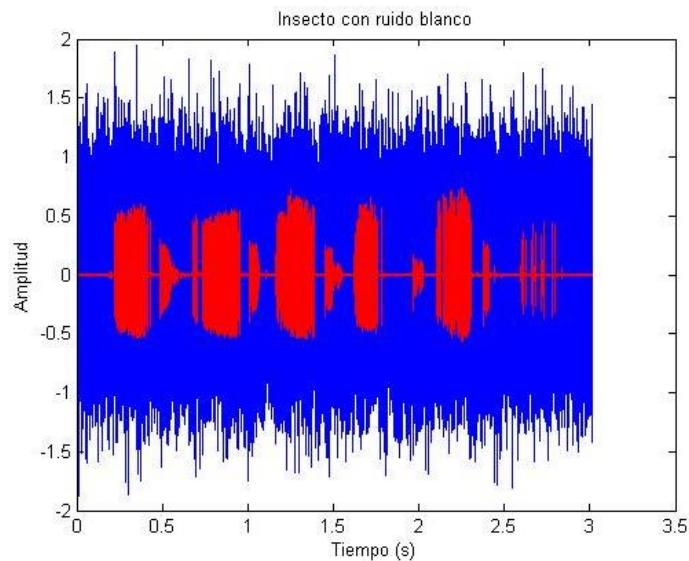


Figura 12 Insecto 1 con ruido blanco gaussiano y un nivel de -15 dB de SNR. En rojo, el insecto original. En azul, el insecto con el ruido añadido.

Podemos observar que al añadirle ruido con tal valor de SNR prácticamente no se es capaz de diferenciar por el oído el sonido del insecto, sin embargo, el sistema implementando NMF sí es capaz de identificar correctamente el insecto en cuestión.

Nos centramos en estos valores de SNR ya que son los más críticos y se hace complicado diferenciar entre ruido e insecto. Obviamente, para valores mayores de SNR se detectan correctamente.

Sin embargo, existen diferentes ficheros de insectos donde al añadir ruido blanco se es incapaz de detectar al usar bases de ruido blanco y se produce error en la detección. Por ejemplo tenemos el insecto 13, donde su forma de onda se muestra en la siguiente figura:

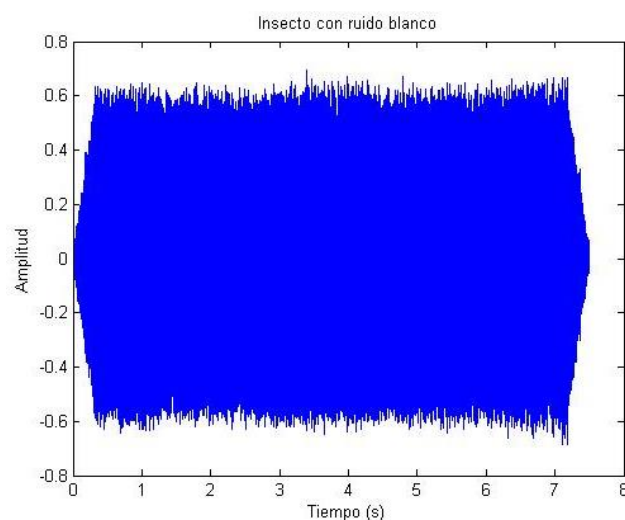


Figura 13 Insecto 13

En la figura superior, vemos la forma de onda de nuestro insecto. Para comprobar la distribución en frecuencia del insecto, se muestra en la figura inferior un espectrograma donde se representa el número de tramas, con una duración de 65 milisegundos, y las frecuencias obtenidas en cada una de ellas.

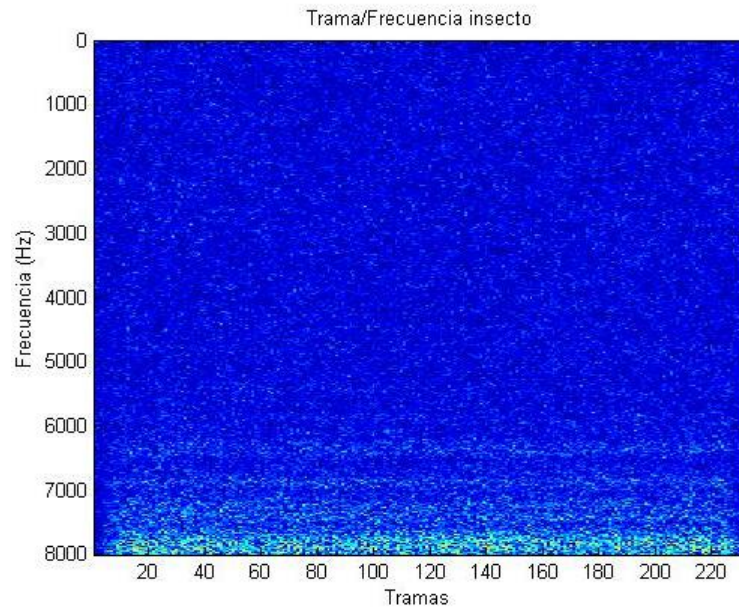


Figura 14 Espectrograma insecto 13

Podemos observar en la figura 13 como el espectro se distribuye prácticamente en altas frecuencias alrededor de los 8 kHz.

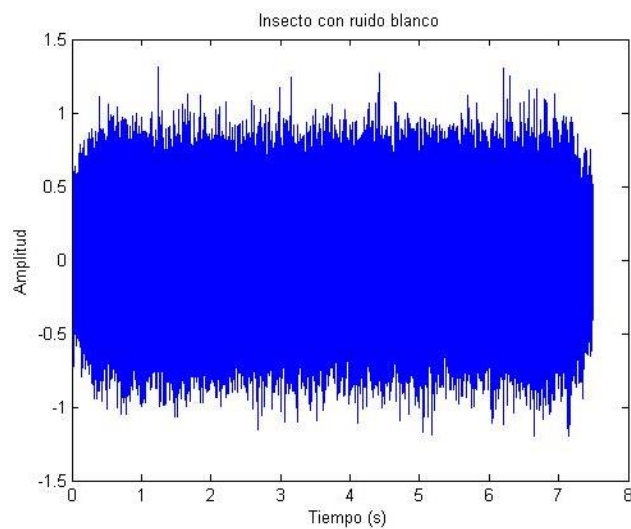


Figura 15 Insecto 13 con ruido blanco y un nivel de SNR de 0 dB

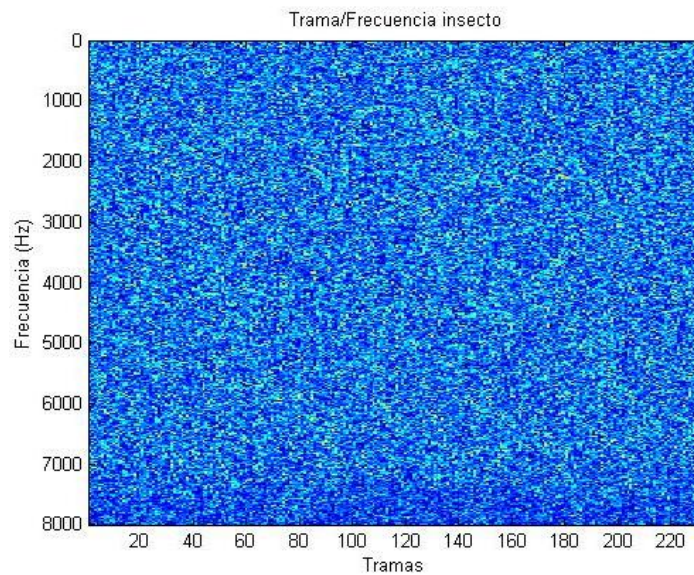


Figura 16 Espectrograma insecto 13 con ruido blanco y un nivel de SNR de 0 dB

En el caso de este insecto, el reconocimiento sin introducir ningún tipo de ruido no es posible al usar bases de ruido blanco. Es debido a que el propio sonido se tomará como tal y se eliminará sin poder reconocer el insecto. Si eliminamos las bases de ruido blanco, al introducir el insecto sin ningún tipo de ruido ya se hace posible, pero al añadirle ruido blanco, vuelve a empeorar el sistema y no se reconoce para ningún valor de SNR. El insecto 17 y 22 tiene un comportamiento similar, por tanto se reserva un apartado para estudiar de manera independiente estos dos insectos.

Terminada esta primera prueba, sabiendo entre qué niveles de SNR se encuentra el funcionamiento de nuestro software, somos capaces de pasar a la siguiente fase para proceder a la adición de ruidos más habituales en el mundo real como se ya se ha comentado.

3.3.2 Resultados añadiendo Noizeus para diferentes valores de los elementos de entrada del software

Conocidos los posibles límites para los niveles de SNR, se ha pasado a realizar pruebas en entornos más comunes. Se ha utilizado ruido de diferentes entornos procedente de la base de datos Noizeus que se han combinado con el ruido del insecto, como por ejemplo, ruido de aeropuerto, de trenes o de restaurantes entre otros. Del mismo modo, se han probado diferentes niveles de SNR (obtenidas de la prueba anterior). Hemos hecho un estudio del funcionamiento del sistema con diferentes bases para los diferentes tipos de ruido definidos en el software (blanco, impulsivo y armónico) y valores de sparseness y smoothness para la activación de éstas.

Los siguientes parámetros, corresponde con los datos de entrada de nuestra función usados. Los valores que configuran el número de bases de los diferentes tipos de ruido a discriminar, se denominan: “rnvalues”, “rpvalues” y “rhvalues”. Siendo bases de ruido blanco, bases de ruido percusivo y bases de ruido armónico respectivamente. Por último, tenemos los valores de sparseness y smoothness (vecKSP y vecKSM), siendo los responsables de la activación de bases más abrupta o aislada (parámetro sparseness) o una serie de bases más consecutiva o suave (smoothness).

rnvalues	rpvalues	rhvalues	vecKSP	vecKSM
----------	----------	----------	--------	--------

Tabla 2 Parámetros de entrada del software

Como ya se ha comentado, se han realizado varias pruebas para ir obteniendo los resultados e ir acotando las posibles mejoras que puedan aplicarse al algoritmo. En una primera ejecución del código se han utilizado como parámetros de entrada, una batería de valores para cada argumento, para una vez obtenidos los resultados poder analizarlos y ver los posibles valores donde el software funciona de manera más eficiente. La batería de valores introducidos para los diferentes argumentos ha sido muy variada. Como por ejemplo:

- Número de bases de ruido: 0, 2, 5 y 10 bases para cada tipo de ruido (impulsivo, armónico y ruido blanco).
- Valores de smoothness y sparseness: 0, 1, 3, 10, 30, y 100 para cada parámetro.
- Valor de relación señal a ruido (SNR): desde 20 a -20 dB con saltos de 5 dB.

Estos valores se han combinado con todas las posibilidades para llevar un estudio exhaustivo del sistema y proceder a acotar al ir obteniendo los resultados. Destacar el gran coste computacional que conlleva, ya que todas las posibilidades de combinación entre todos los valores es muy elevada (2300 identificaciones aproximadamente por cada valor de SNR).

Como ruido se ha utilizado la base de datos Noizeus con 8 tipos de ruido diferentes. Se incluyen ruido de aeropuerto, de conversaciones, de tren y coche, de calle, de estación de tren, de restaurante y de sala de exposiciones. Para cada tipo de ruido, se han utilizado los valores mencionados anteriormente.

De manera generalizada, el sistema funciona correctamente cuando usamos bases para eliminar el ruido blanco y se ha comprobado la mejora del sistema cuando se utilizan el mismo número de bases para todos los tipos de ruido. Como ejemplo, se muestra una tabla donde se plasman algunos de los valores de los diferentes parámetros donde no se realiza un reconocimiento exitoso.

rnvalues	rpvalues	rhvalues	vecKSP	vecKSM
0	10	10	30	1
0	10	10	30	3
0	10	10	30	10
0	10	10	30	30
0	10	10	30	100
0	10	10	100	0
0	10	10	100	1
0	10	10	100	3
0	10	10	100	10
0	10	10	100	30
0	10	10	100	100

Tabla 3 Errores en identificación insecto 2 con ruido de aeropuerto y nivel de SNR de 10 dB

En la tabla superior, se pueden ver los diferentes valores que se le han introducido en los que se han producido error en la identificación del insecto. Los diferentes valores marcados en la tabla, corresponde con los datos de entrada usados de nuestra función como se ha comentado anteriormente. Se aprecia que todos los valores producen un reconocimiento incorrecto en el caso de no utilizar valores en el primer término (bases de ruido blanco). Se tiene una situación similar para todos los tipos de ruido e insectos, donde es necesario el uso de estas bases para lograr una correcta identificación. En los siguientes gráficos, se excluyen los insectos 13, 17 y 22 para su estudio aparte.

3.3.2.1 Ruido aeropuerto

El primero de los ruidos utilizados, ha sido ruido producido en el interior de un aeropuerto. Como vemos en la gráfica, con este tipo de ruido se logra detectar con un alto porcentaje de éxito desde los 20 dB hasta 0 dB. Cuando operamos con valores negativos de SNR comenzamos a tener peores porcentajes de éxito teniendo un 30% para un nivel de – 10 dB.

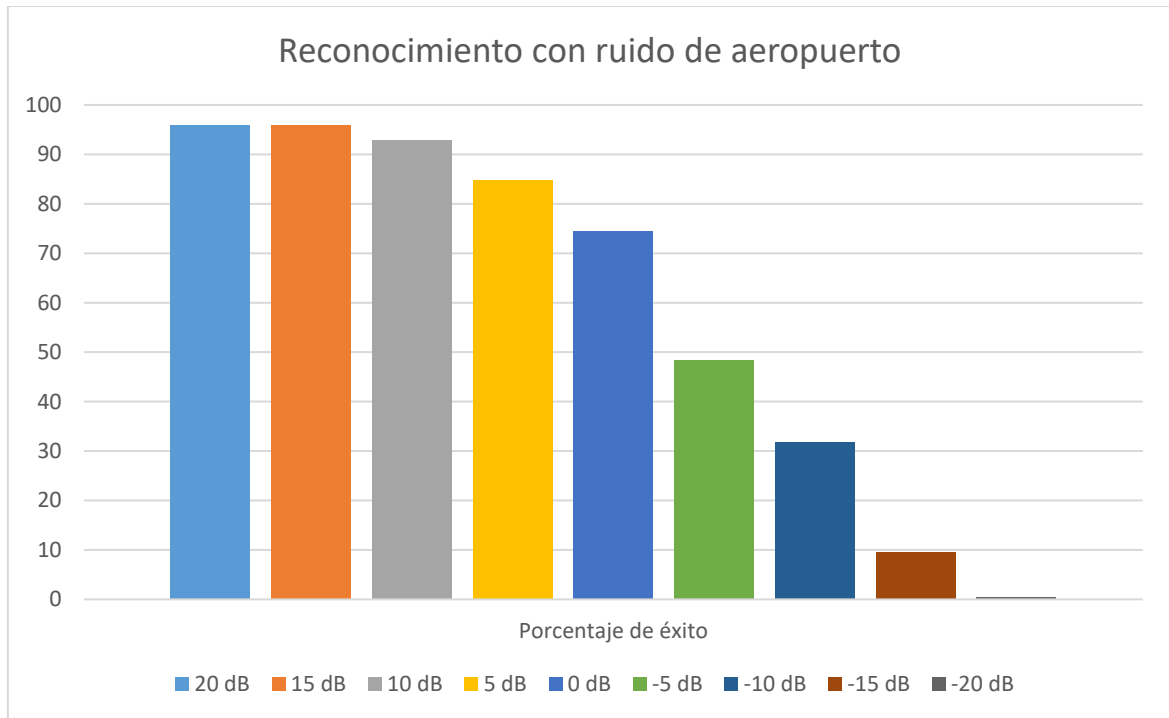


Figura 17 Porcentaje de éxito en detección para todos los insectos con ruido de aeropuerto. 1150 operaciones por cada nivel de SNR. Configuración de 5 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

Si comparamos el uso de 5 bases de caracterización con el uso de 3 bases, usando ambos un valor de sparseness y smoothness igual a 1 (modelo óptimo como se verá más adelante), vemos que se produce una mejora que aumenta conforme disminuimos los niveles de SNR, hasta llegar a los valores negativos donde ya se tiene un comportamiento muy similar en ambos casos.

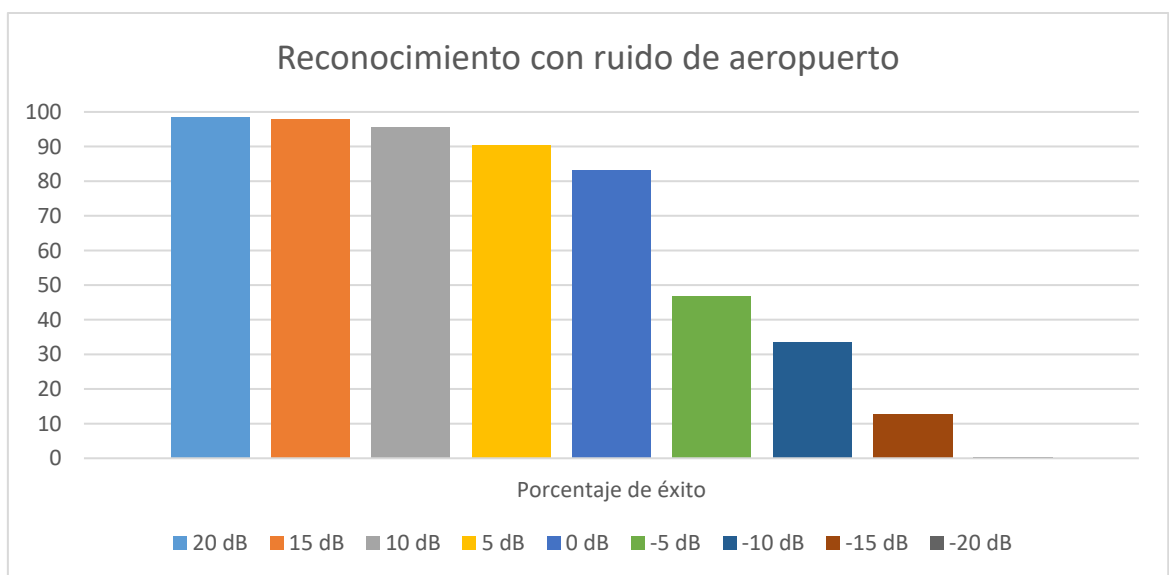


Figura 18 Porcentaje de éxito en detección para todos los insectos con ruido de aeropuerto. 1150 operaciones por cada nivel de SNR. Configuración de 3 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

La forma de onda y espectrograma del ruido en cuestión es:

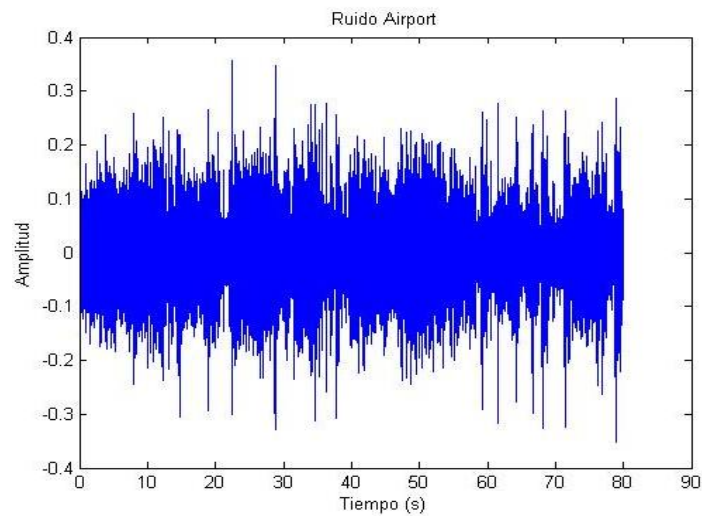


Figura 19 Ruido de aeropuerto

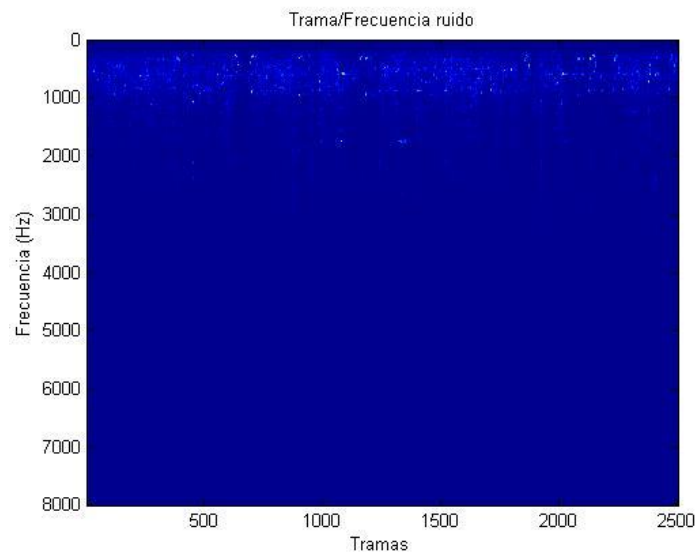


Figura 20 Espectrograma ruido de aeropuerto

Con este tipo de ruido, en general para SNR altas (20 - 0 dB) se tiene una alta probabilidad de reconocer el insecto dado. A medida que comenzamos a disminuir la SNR se comienzan a apreciar los problemas

3.3.2.2 Ruido de conversación

En este tipo de ruido, tenemos un comportamiento similar al anterior. Es necesario el usar bases de ruido blanco para comenzar a identificar correctamente el insecto.

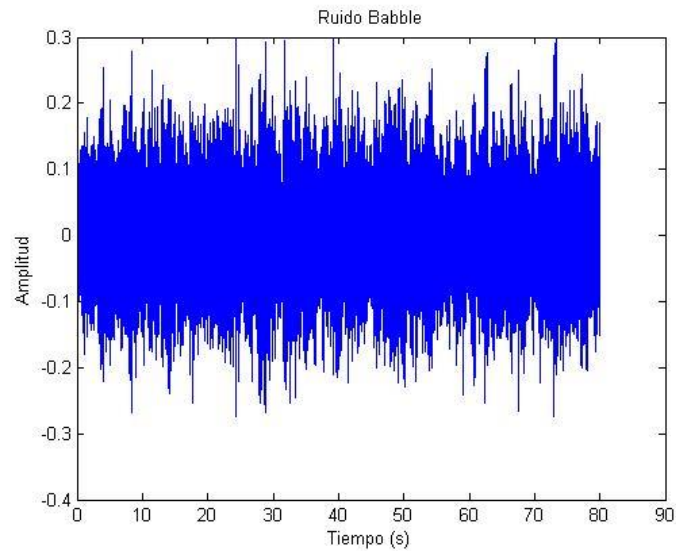


Figura 21 Ruido de conversación

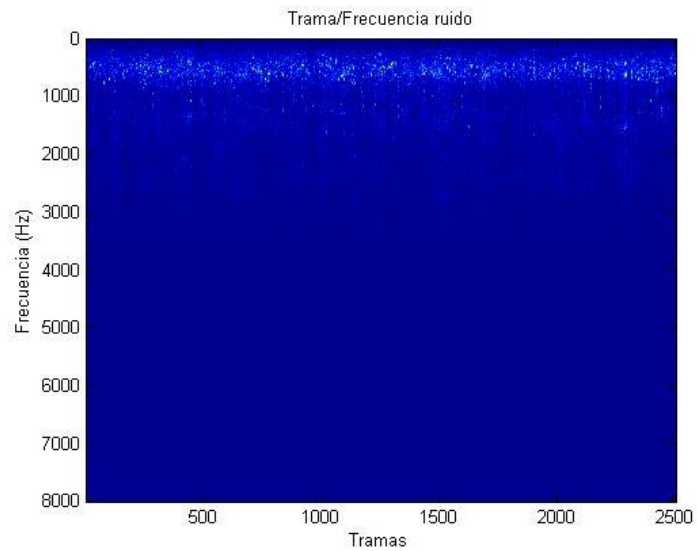


Figura 22 Espectrograma ruido de conversación

Tiene un comportamiento muy similar al ruido anterior entre los 20 y 0 dB. Si seguimos disminuyendo la SNR. A continuación se muestra una figura donde se muestran los porcentajes de éxito para cada nivel de SNR con 1150 operaciones por cada una de ellas.

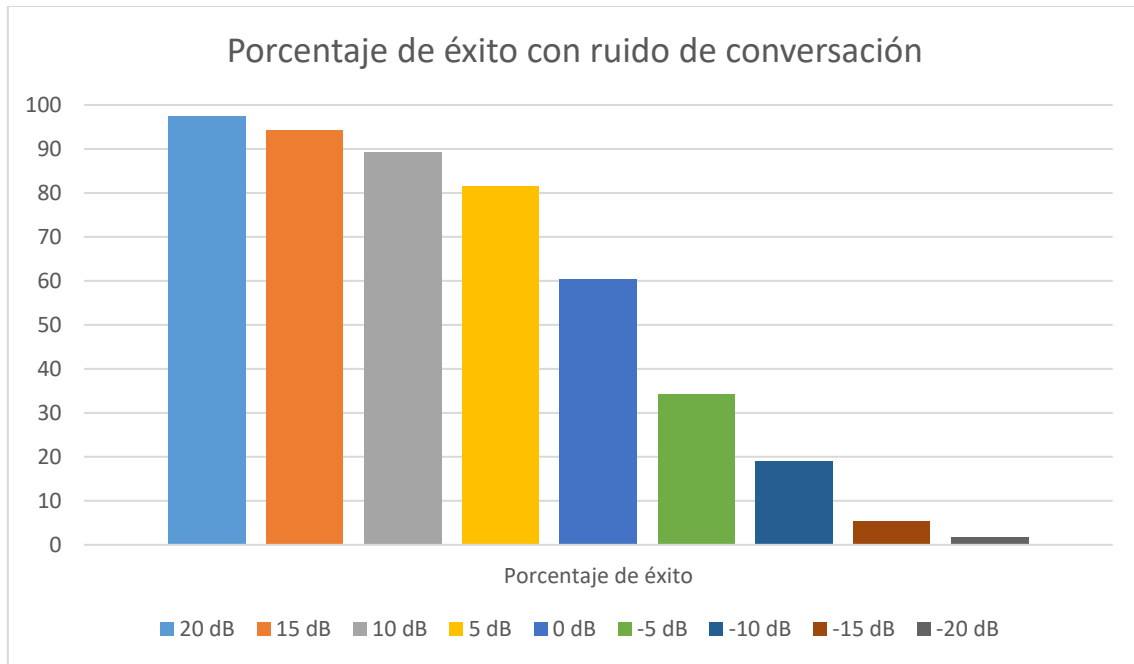


Figura 23 Porcentaje de éxito en detección para todos los insectos con ruido de conversación. 1150 operaciones por cada nivel de SNR. Configuración de 5 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

De nuevo tenemos el mismo caso anterior, donde se mejora con la configuración de 3 bases de caracterización y valores de sparseness y smoothness igual a 1.

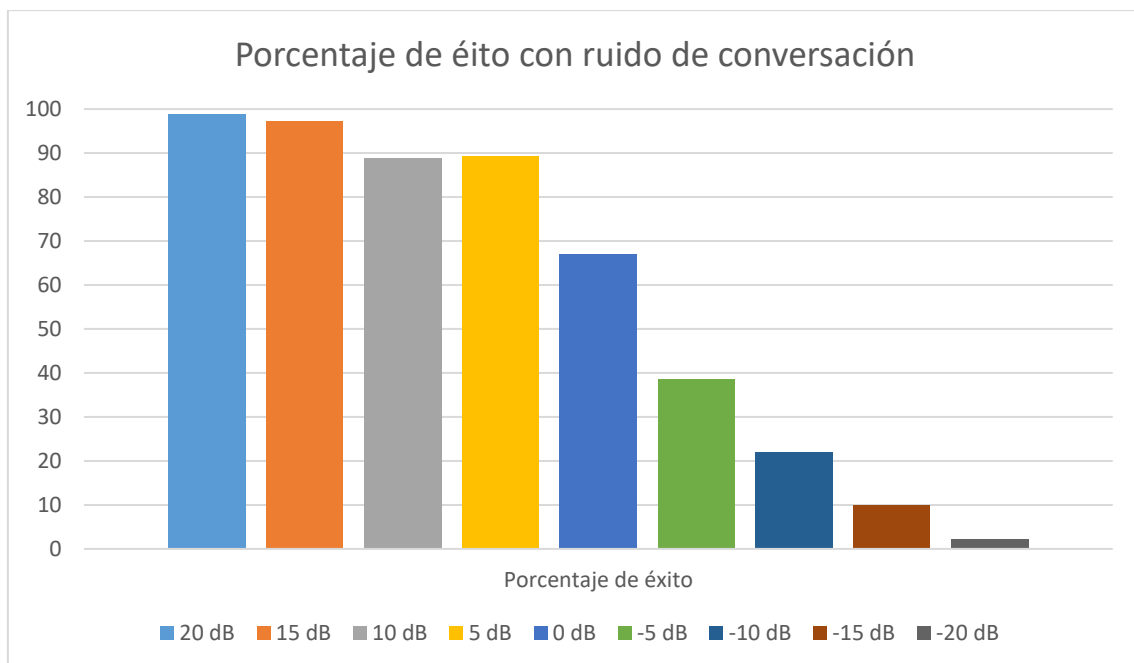


Figura 24 Porcentaje de éxito en detección para todos los insectos con ruido de conversación. 1150 operaciones por cada nivel de SNR. Configuración de 3 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

Al disminuir el nivel de SNR (-5 dB), ya tenemos un porcentaje de éxito menos del 40 % el cual sigue disminuyendo conforme disminuimos el nivel de SNR.

3.3.2.3 Ruido de coche

Como vemos el comportamiento con casi todos los tipos de ruido es muy similar y mantiene una misma tendencia, donde al llegar a un nivel de 0 dB se comienza a degradar el porcentaje de éxito de nuestro reconocimiento. En este caso, vemos como se mantiene por encima del 50% con un nivel de -5 dB y se mantiene por encima del 30% para -10 dB en el caso de utilizar 5 bases de caracterización.

Como se ha visto en los ruidos anteriores, el comportamiento sigue siendo el mismo y se confirma la mejora con el uso de 3 bases para caracterizar los tipos de ruido y los insectos.

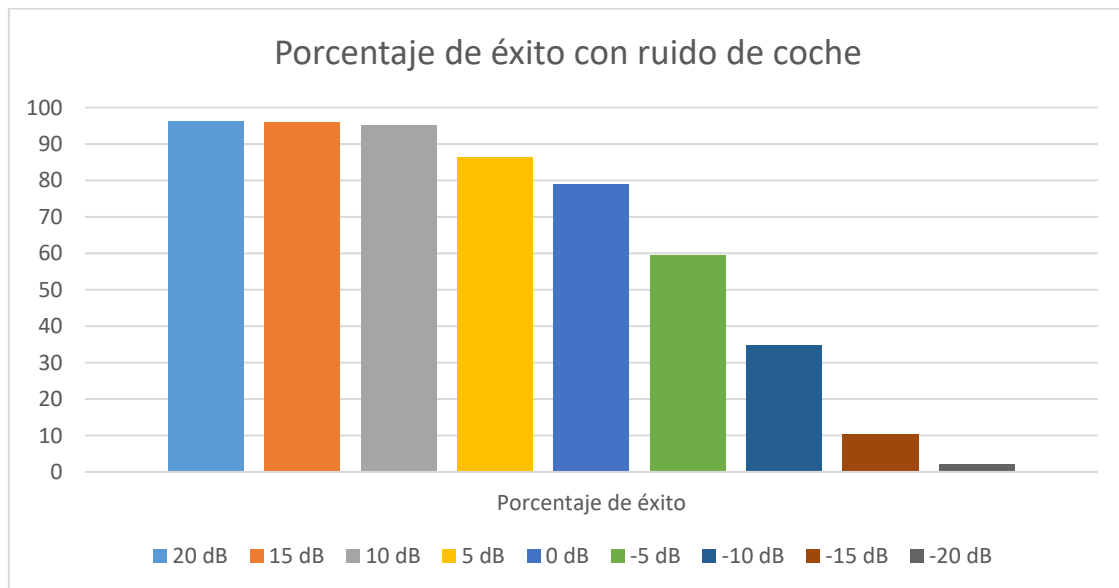


Figura 25 Porcentaje de éxito en detección para todos los insectos con ruido de coche. 1150 operaciones por cada nivel de SNR. Configuración de 5 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

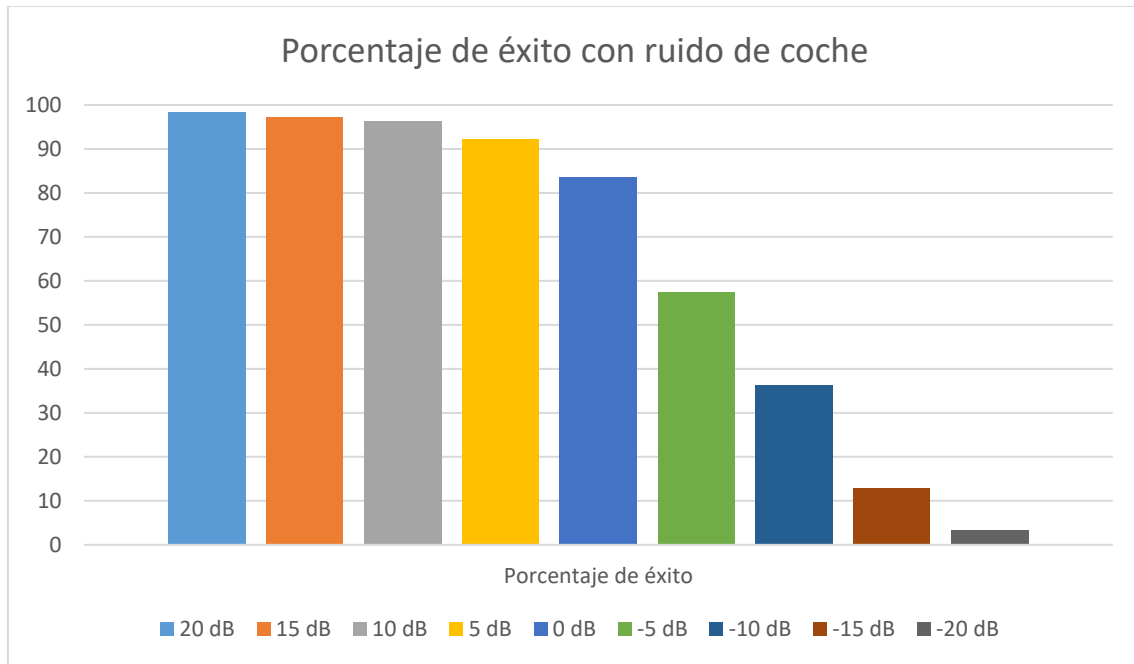


Figura 26 Porcentaje de éxito en detección para todos los insectos con ruido de coche. 1150 operaciones por cada nivel de SNR. Configuración de 3 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

La forma de onda y el espectrograma característico de este tipo de ruido se muestra a continuación.

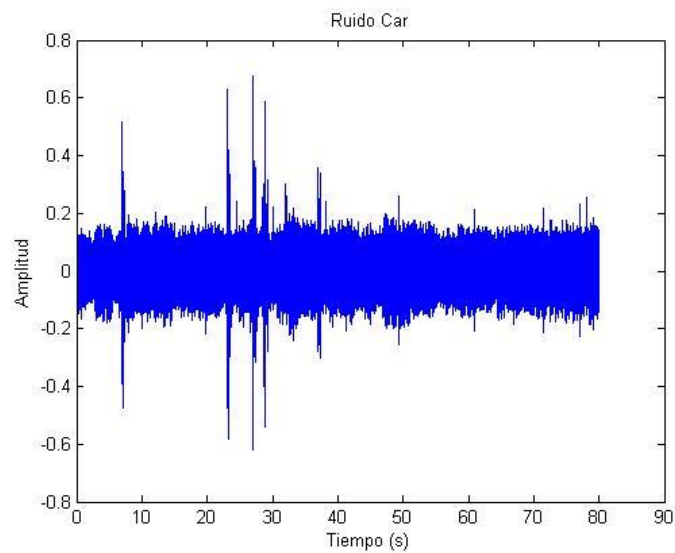


Figura 27 Ruido de coche

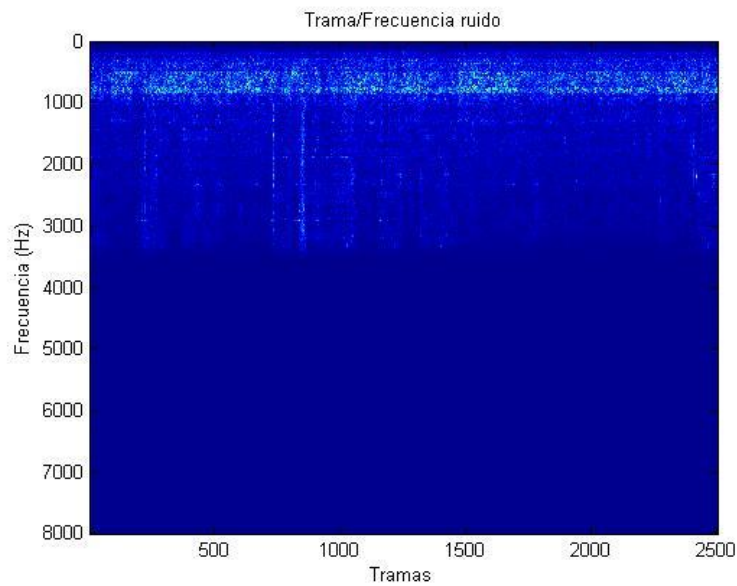


Figura 28 Espectrograma ruido de coche

3.3.2.4 Ruido de sala de exposiciones

Usando este ruido, se tiene un comportamiento muy similar al obtenido en el caso anterior con valores de SNR positivos, sin embargo para valores negativos se degrada más rápidamente el porcentaje de éxito.

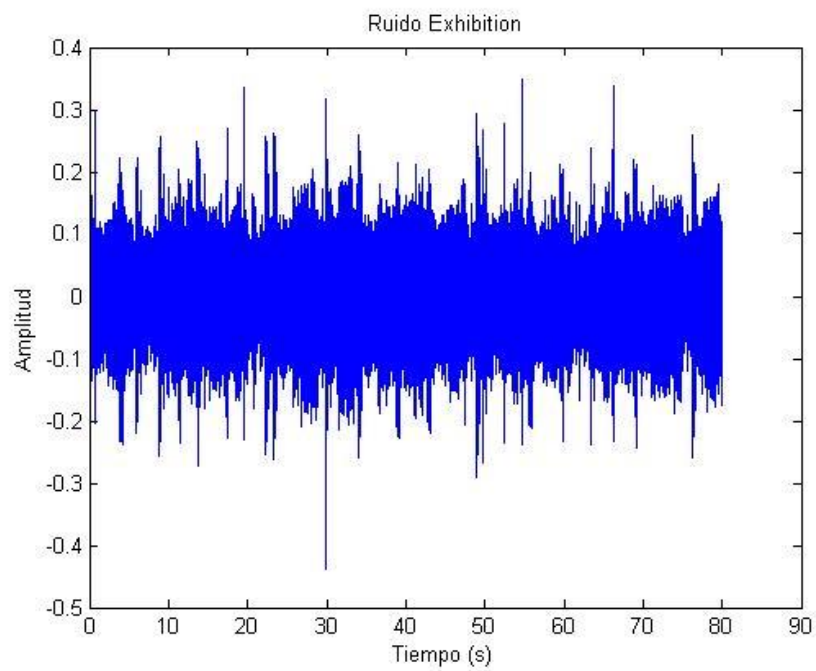


Figura 29 Ruido de sala de exposiciones

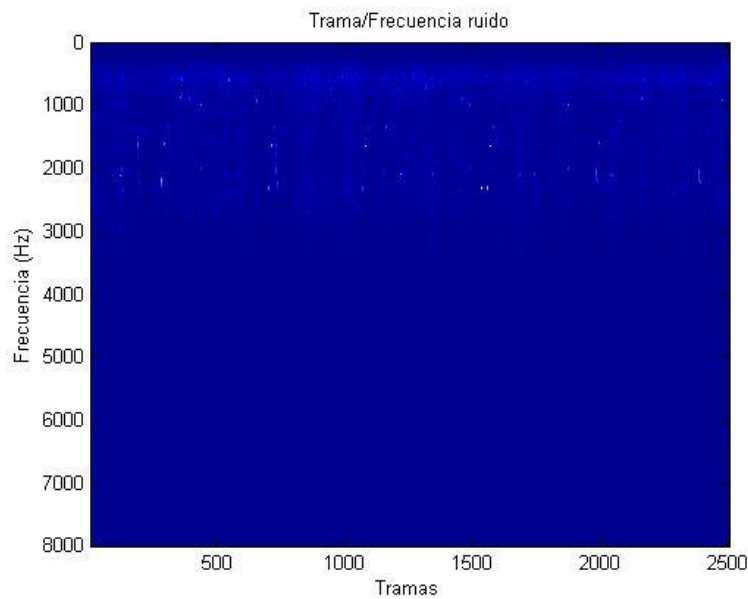


Figura 30 Espectrograma ruido de sala de exposiciones

Tenemos un comportamiento aceptable hasta casi -5 dB donde ya, al pasar a un nivel de SNR de -10 dB tenemos una disminución de un casi 30% con respecto el nivel anterior para el caso de utilizar 5 bases de caracterización.

Si nos centramos en el uso de 3 bases, donde se plasma una mejora para todos los niveles de SNR mayores de -10 dB, vemos un mantenimiento del porcentaje de éxito para los valores elevados de SNR (20, 15 y 10 dB) y se tiene un mayor éxito para valores menores de SNR en comparación la utilización de 5 bases de caracterización.

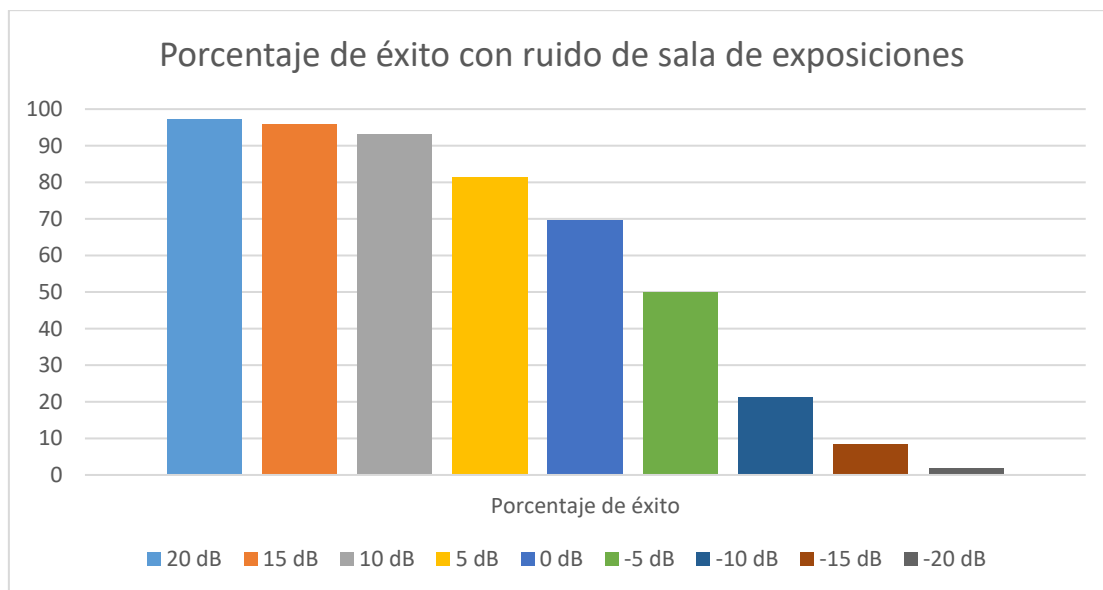


Figura 31 Porcentaje de éxito en detección para todos los insectos con ruido de sala de exposiciones. 1150 operaciones por cada nivel de SNR. Configuración de 5 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

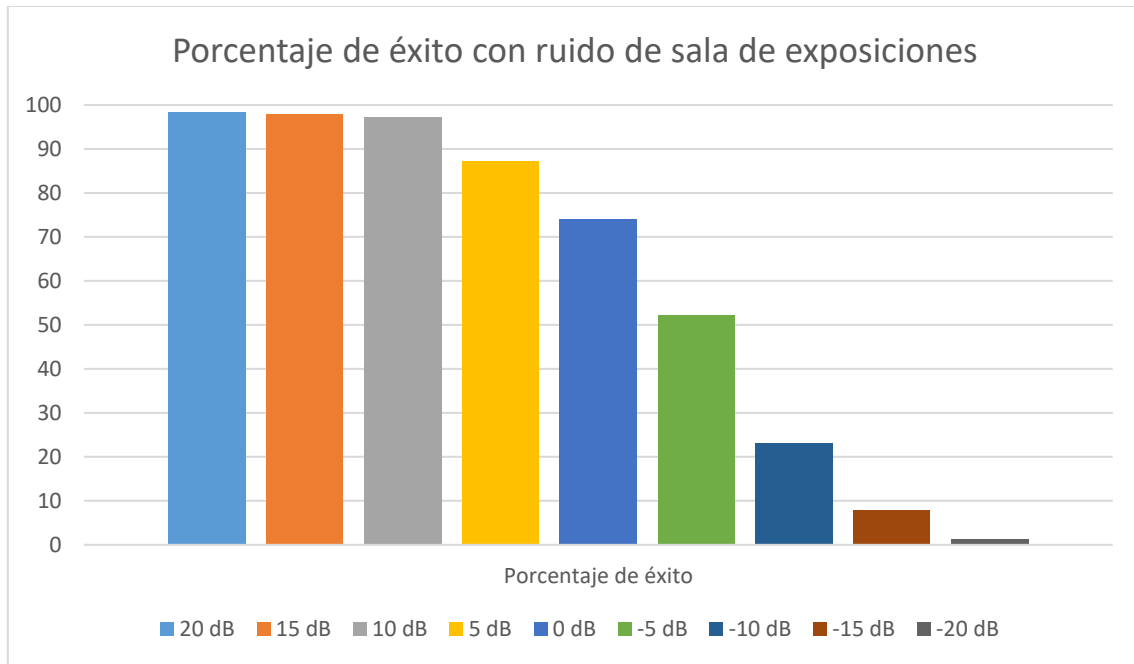


Figura 32 Porcentaje de éxito en detección para todos los insectos con ruido de sala de exposiciones. 1150 operaciones por cada nivel de SNR. Configuración de 3 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

3.3.2.5 Ruido de restaurante

Con niveles de altos de SNR no tenemos problema en la detección. Se tiene un comportamiento similar como en los casos anteriores.

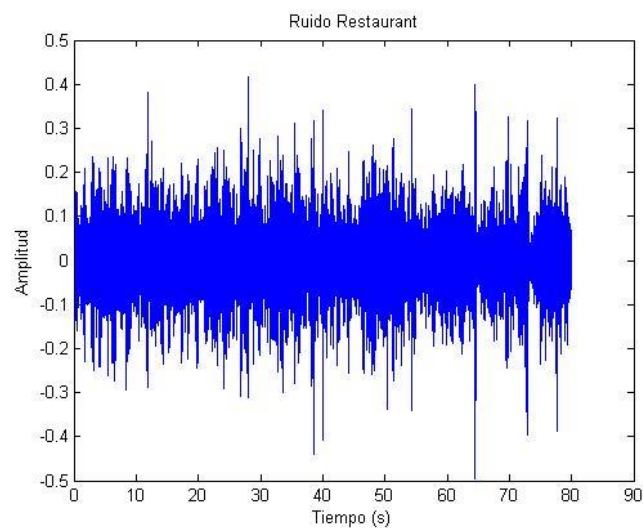


Figura 33 Ruido de restaurante

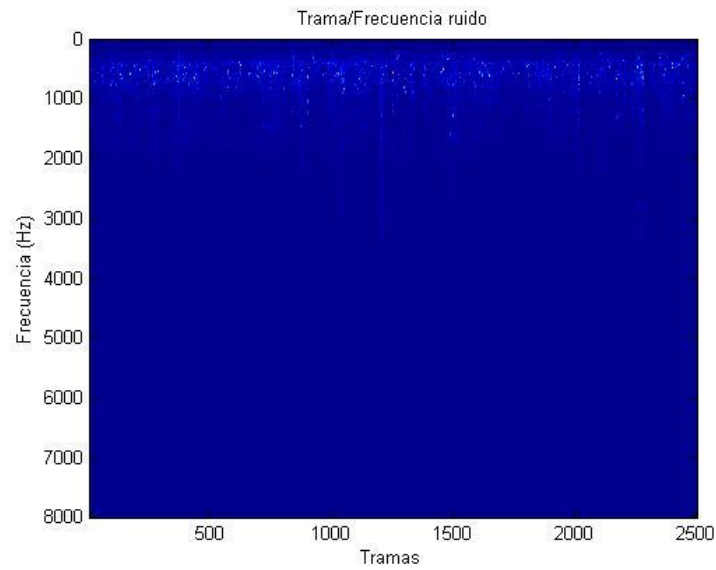


Figura 34 Espectrograma ruido de restaurante

Como vemos en la gráfica, al igual que en los casos anteriores, se tiene un buen comportamiento hasta la frontera del nivel de SNR de 0 dB. A partir de este punto comienza a disminuir el porcentaje de éxito de manera más abrupta.

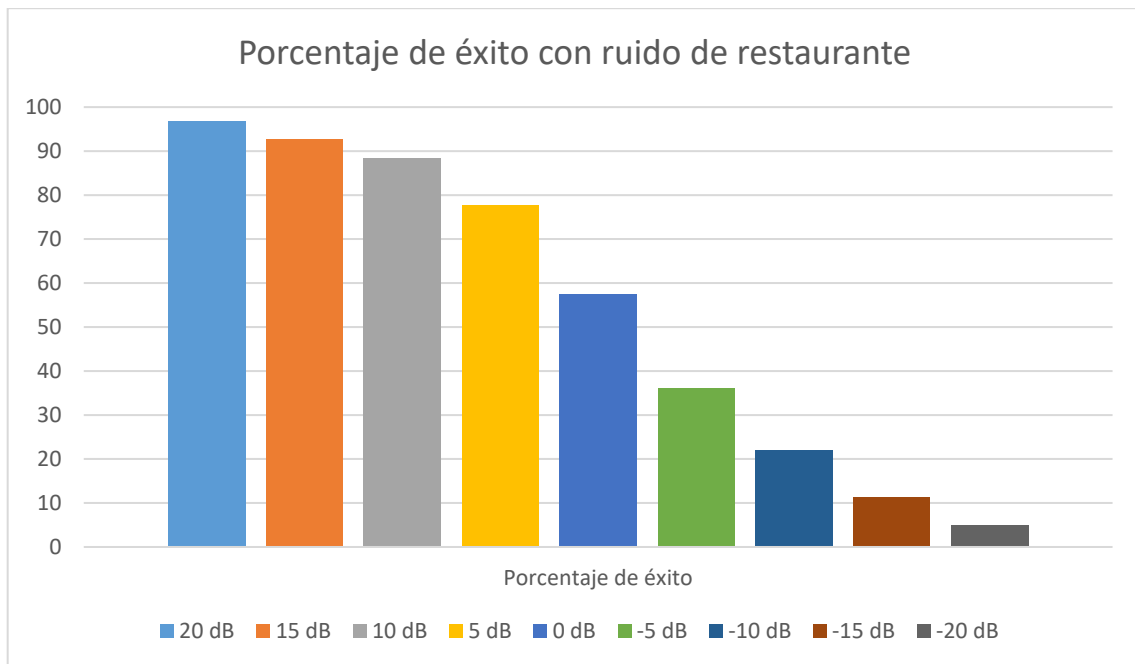


Figura 35 Porcentaje de éxito en detección para todos los insectos con ruido de restaurante. 1150 operaciones por cada nivel de SNR. Configuración de 5 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

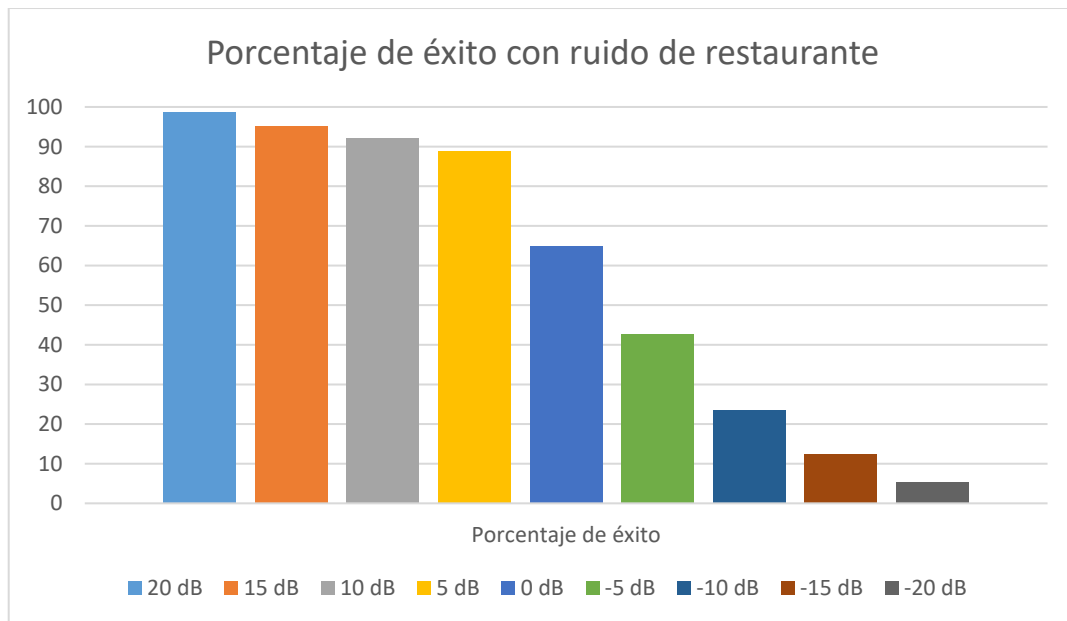


Figura 36 Porcentaje de éxito en detección para todos los insectos con ruido de restaurante. 1150 operaciones por cada nivel de SNR. Configuración de 3 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

De nuevo, seguimos teniendo una mejora con el uso de 3 bases de caracterización en comparación a la utilización de 5 bases.

3.3.2.6 Ruido de estación de tren

Para el ruido de estación, con niveles altos de SNR no se suelen tener problemas a la hora de identificar. Aunque depende del insecto estudiado, por lo general se cumple para casi todos. Como veremos en el gráfico, los porcentajes de éxito con este tipo de ruido son similares a los anteriores.

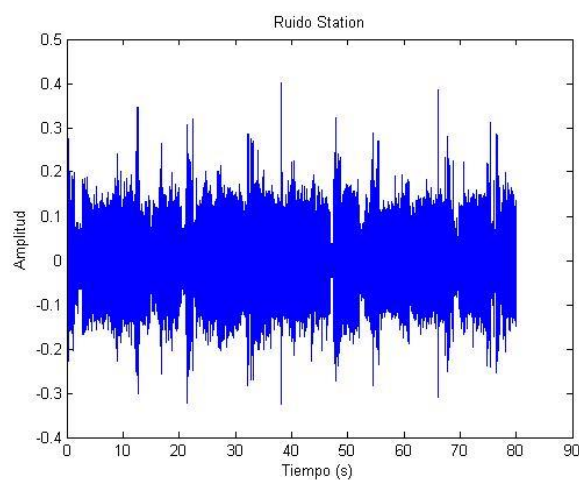


Figura 37 Ruido de estación

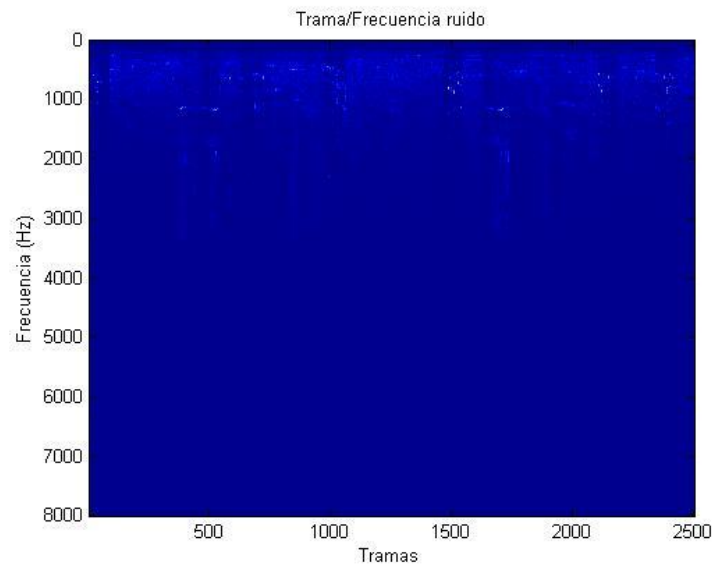


Figura 38 Espectrograma ruido de estación de tren

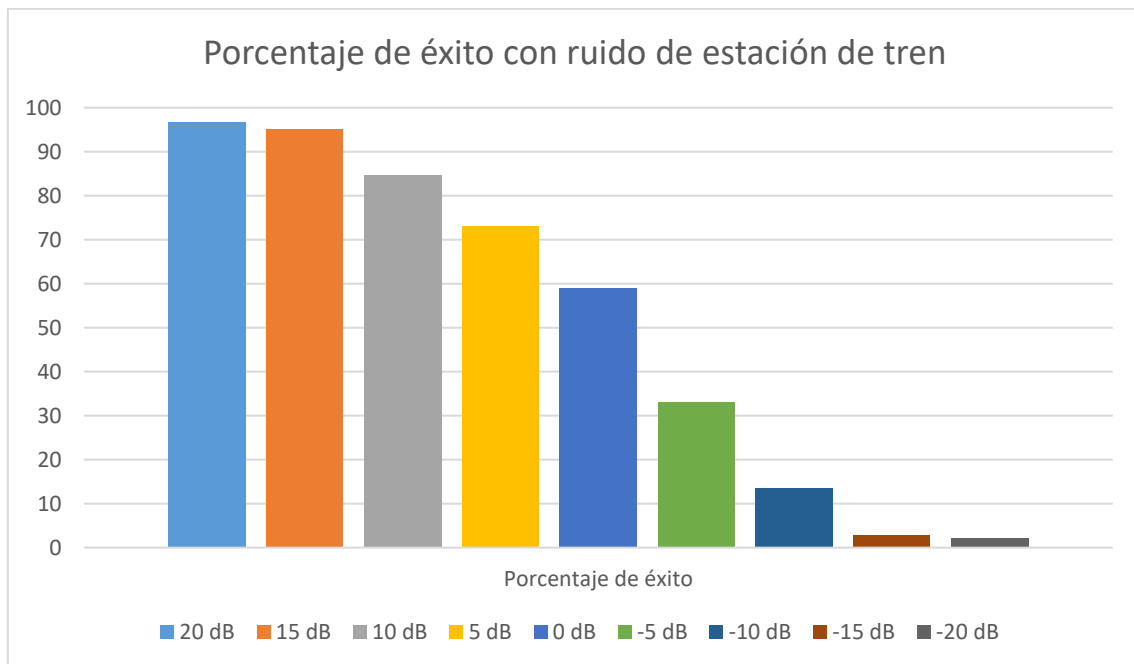


Figura 39 Porcentaje de éxito en detección para todos los insectos con ruido de estación de tren. 1150 operaciones por cada nivel de SNR. Configuración de 5 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

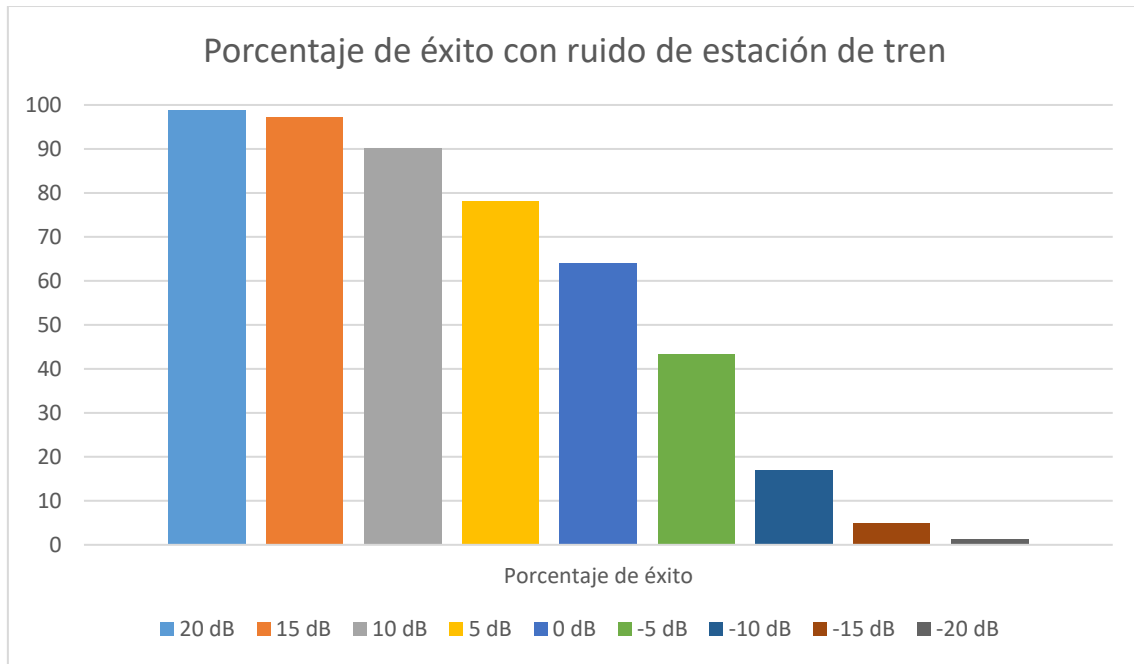


Figura 40 Porcentaje de éxito en detección para todos los insectos con ruido de estación de tren. 1150 operaciones por cada nivel de SNR. Configuración de 3 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

Podemos apreciar el mismo comportamiento comentado anteriormente. Destacar un descenso menos abrupto del porcentaje de éxito a media que disminuimos el nivel de SNR para el uso de 3 bases de caracterización.

3.3.2.7 Ruido de calle

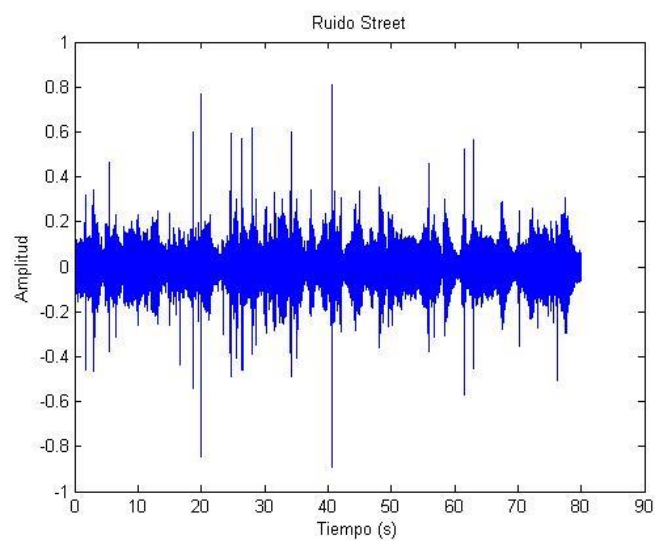


Figura 41 Ruido de calle

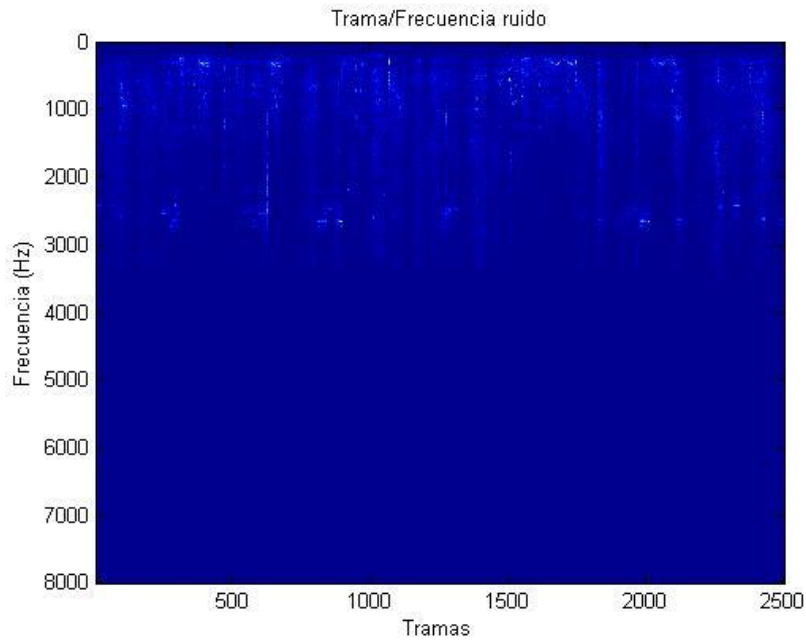


Figura 42 Espectrograma ruido de calle

Para este ruido, de nuevo tenemos el mismo comportamiento descrito en los apartados anteriores. Con la utilización de 3 bases tenemos una mejor identificación y para todos los niveles de SNR en comparación del uso de 5 bases de caracterización.

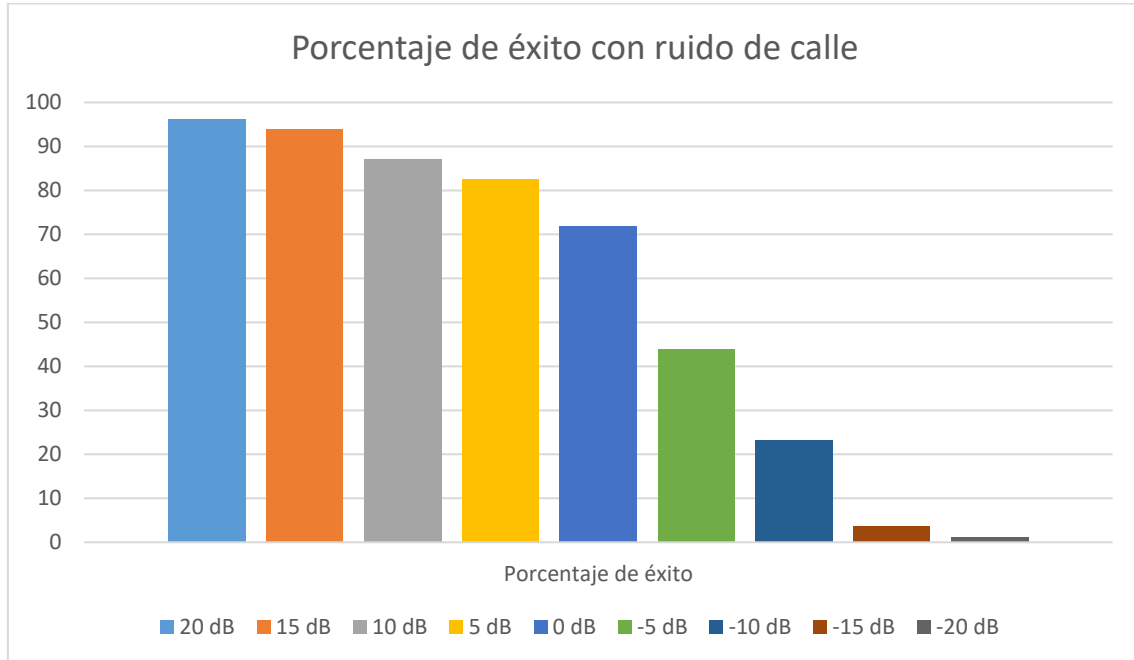


Figura 43 Porcentaje de éxito en detección para todos los insectos con ruido de calle. 1150 operaciones por cada nivel de SNR. Configuración de 5 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

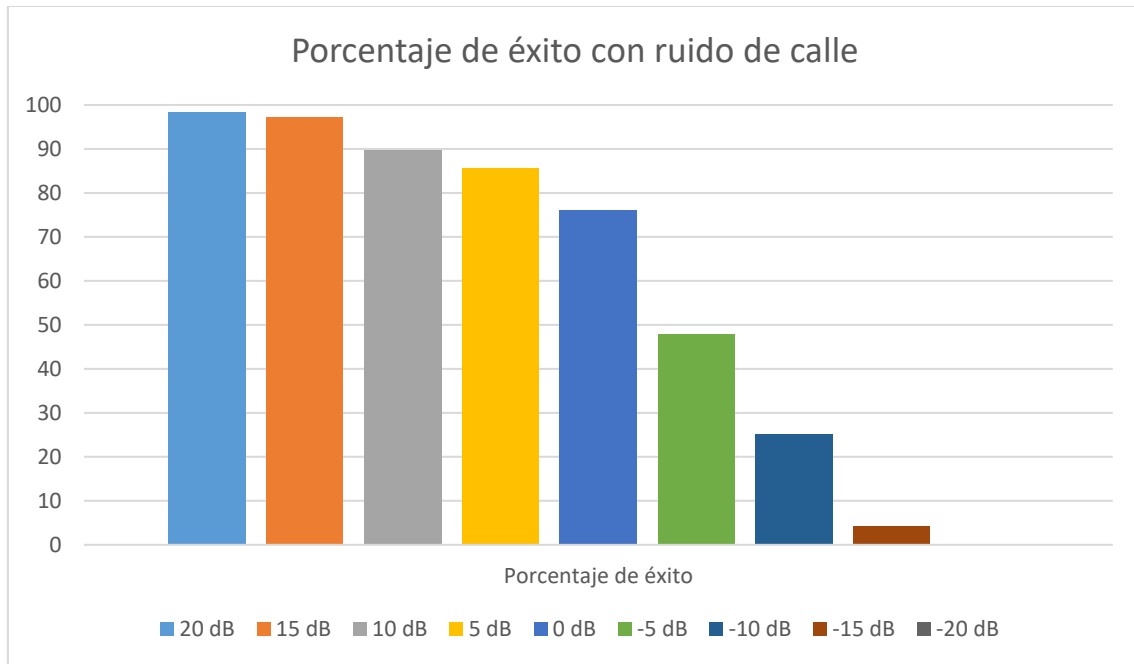


Figura 44 Porcentaje de éxito en detección para todos los insectos con ruido de calle. 1150 operaciones por cada nivel de SNR. Configuración de 3 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

3.3.2.8 Ruido de tren

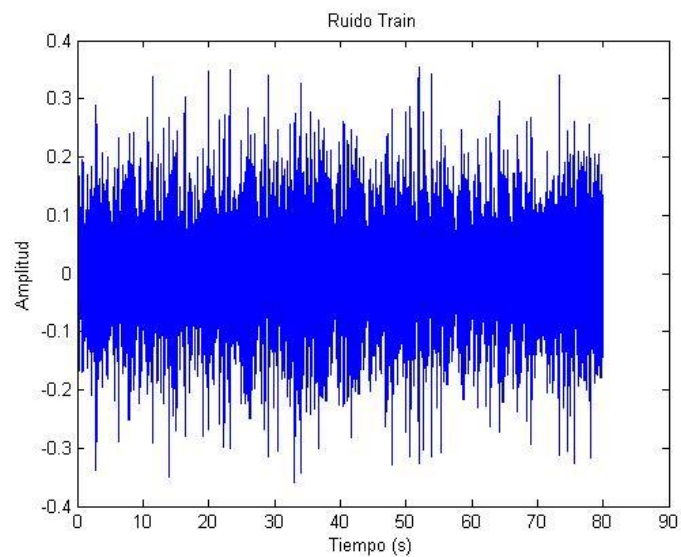


Figura 45 Ruido de tren

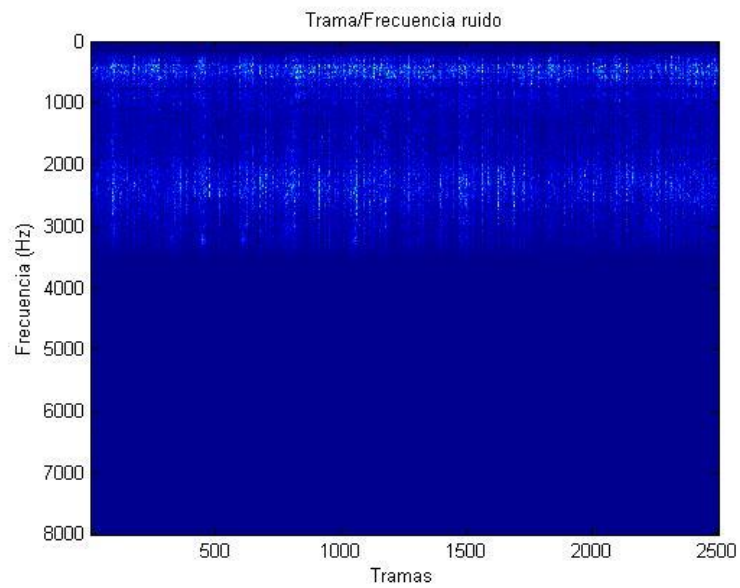


Figura 46 Espectrograma ruido Train

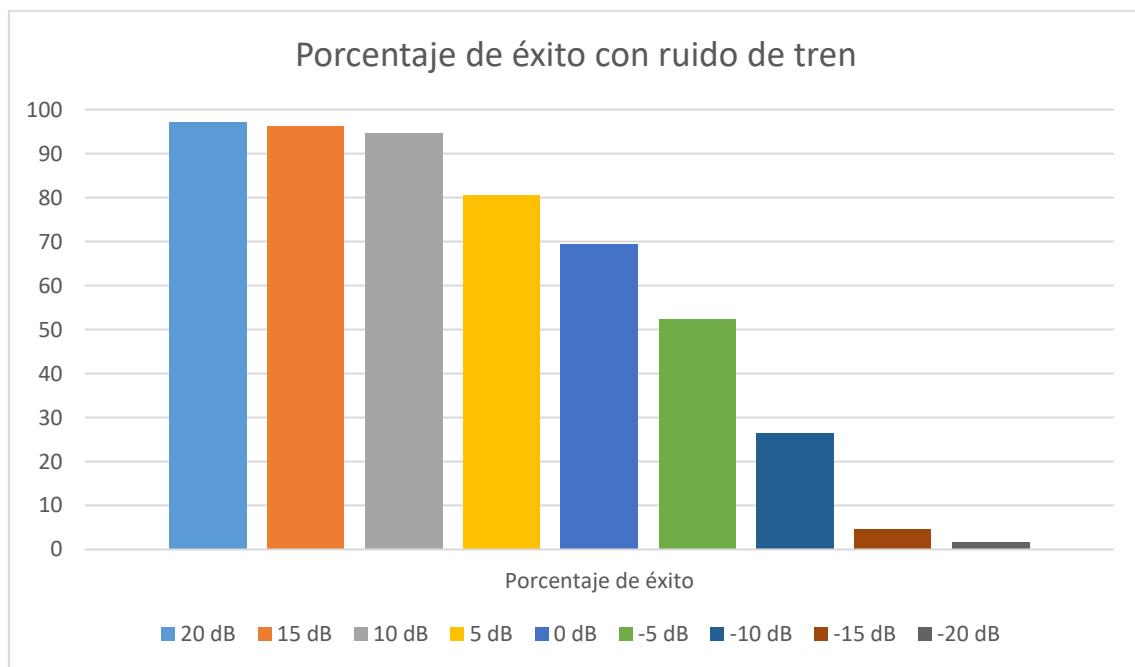


Figura 47 Porcentaje de éxito en detección para todos los insectos con ruido de tren. 1150 operaciones por cada nivel de SNR. Configuración de 5 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

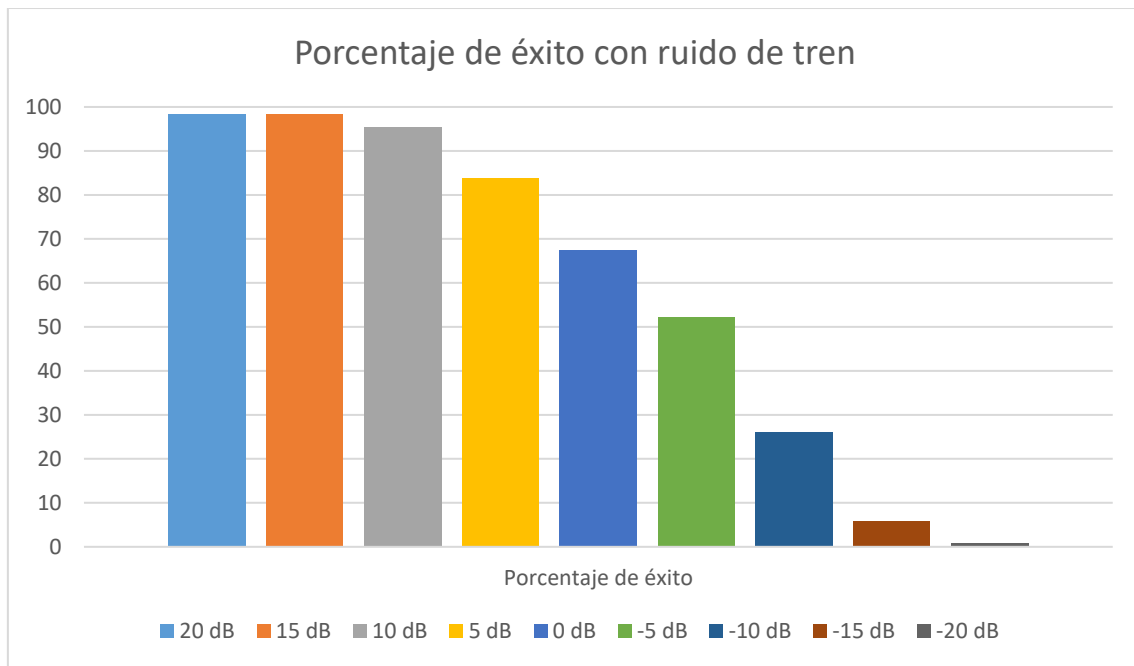


Figura 48 Porcentaje de éxito en detección para todos los insectos con ruido de tren. 1150 operaciones por cada nivel de SNR. Configuración de 3 bases de caracterización de ruido y de insecto y con valor 1 para sparseness y smoothness.

3.3.2.9 Primeras conclusiones sobre el uso de diferentes bases en el reconocimiento de insectos

Se puede ver en esta prueba con los diferentes tipos de ruido y diferentes niveles de SNR que el comportamiento en cuanto a porcentajes de éxito de reconocimiento de insectos, es muy similar para todos ellos. Podemos apreciar un buen comportamiento para niveles positivos de SNR. Cuando comenzamos a tener niveles negativos, se produce una reducción importante en la tasa de éxito, sin embargo se produce a partir de niveles de -10 dB, donde para -5 dB se ronda el 50 % de éxito para casi todos los casos. Notamos también la mejora del uso de 3 bases de caracterización sobre 5 para todos los tipos de ruido.

3.3.3 Resultados añadiendo Noizeus con 5 bases de caracterización

En esta prueba, se han aplicado las conclusiones obtenidas en el caso anterior para ir acotando el número de valores de cada elemento. En este caso, se ha procedido a mantener fijas las bases que caracterizan cada tipo de ruido con un valor idéntico en ellas, ya que se había visto en el caso anterior que los mejores resultados se daban cuando las bases coincidían. Se han fijado en 5 bases para cada tipo de ruido y para la definición del insecto se ha usado el mismo número de bases.

Para realizar la prueba, se han variado todos los valores de smoothness y sparseness (suavidad y escasez) con valores como 0, 1, 3, 10, 30 y 100.

En este caso, se han realizado un total de 1288 operaciones por cada combinación posible. Para cada una se ha llevado a cabo la generación de un nuevo diccionario, compuesto por las distintas bases que caracterizan cada insecto, entrenado con la base de datos de sonidos de insectos.

A continuación se muestra un gráfico donde se observan los resultados con todas las combinaciones de sparseness y smoothness, para 5 bases de caracterización de todos los tipos de ruido e insecto.

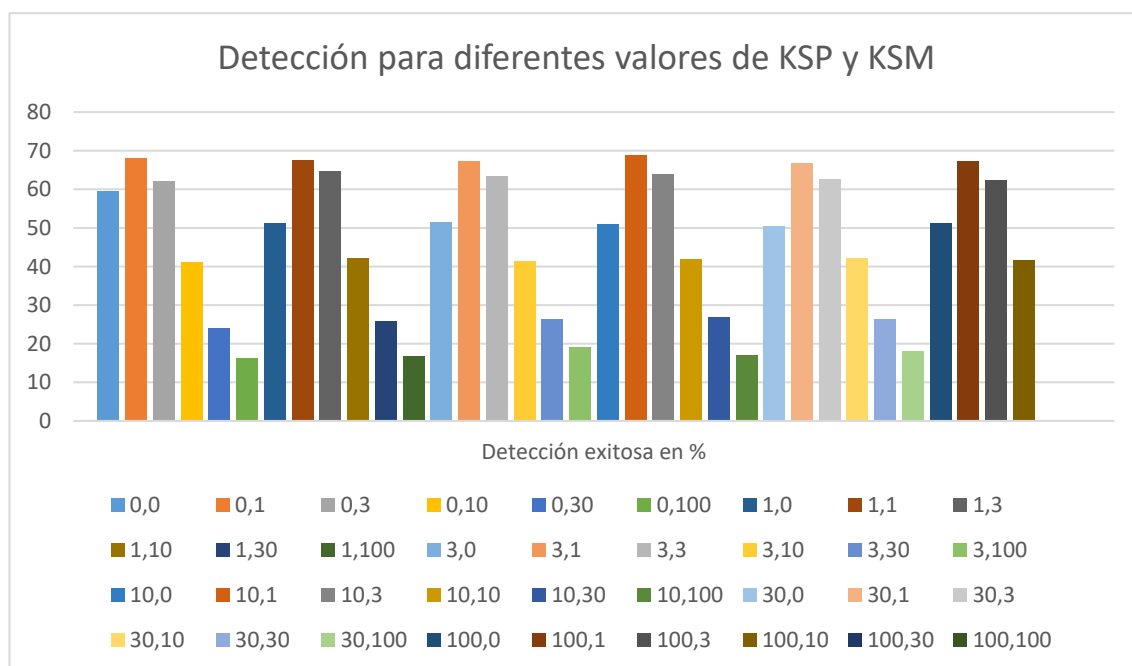


Figura 49 Gráfico de porcentaje de éxito en detección con 5 bases para todos los tipos de ruido y caracterización del insecto. Los valores mostrados corresponden la combinación de sparseness y smoothness respectivamente. Se toman en cuenta todos los tipos de ruido y los valores de SNR (20 hasta -10dB). El número total de operaciones por cada combinación es 1288.

Como se puede observar en el gráfico, los valores de sparseness no afectan al reconocimiento, ya que todos los conjuntos de valores siguen la misma forma mientras se varía el valor de smoothness.

Para el ruido de aeropuerto, se sigue un patrón para casi todos los insectos donde para los valores de sparseness y smoothness bajos funciona el sistema. El comportamiento respecto a la identificación es muy similar en casi todos, donde se realiza una identificación positiva con los mismos valores de suavidad y escasez. En concreto se ha visto que para casi todos los tipos de ruido el software funciona al utilizar valores bajos como 0, 1 y 3 de smoothness. Nos vamos a centrar en estos valores ya que representan un porcentaje de

éxito por encima del 50% y se van a mostrar el porcentaje de éxito con respecto al nivel de SNR.

Como vemos, la mejor configuración se nos muestra para valor de smoothness y sparseness igual a 1. Así mismo vemos que para los valores 0 y 3 tenemos una notable diferencia, donde para niveles de SNR menores si tiene una mejor respuesta para la configuración smoothness y sparseness igual a 3, respecto de 0.

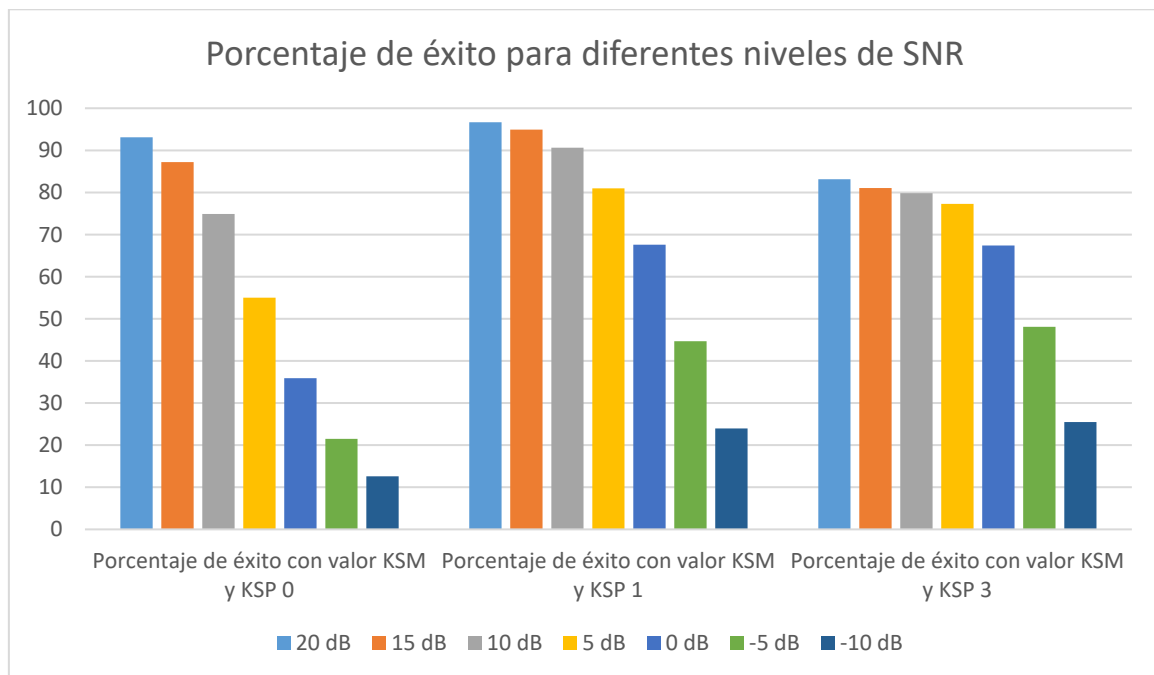


Figura 50 Gráfico de porcentaje de éxito en detección con 5 bases para todos los tipos de ruido y caracterización del insecto. Los valores mostrados al éxito en la detección de insectos para cada nivel de SNR. Se toman en cuenta todos los tipos de ruido. El número total de operaciones por nivel de SNR es 1056.

Existen ciertos tipos de insectos, donde se produce lo contrario. El valor que mejor se adecúa a la identificación son valores elevados de smoothness como por ejemplo en el insecto 12 y 14.

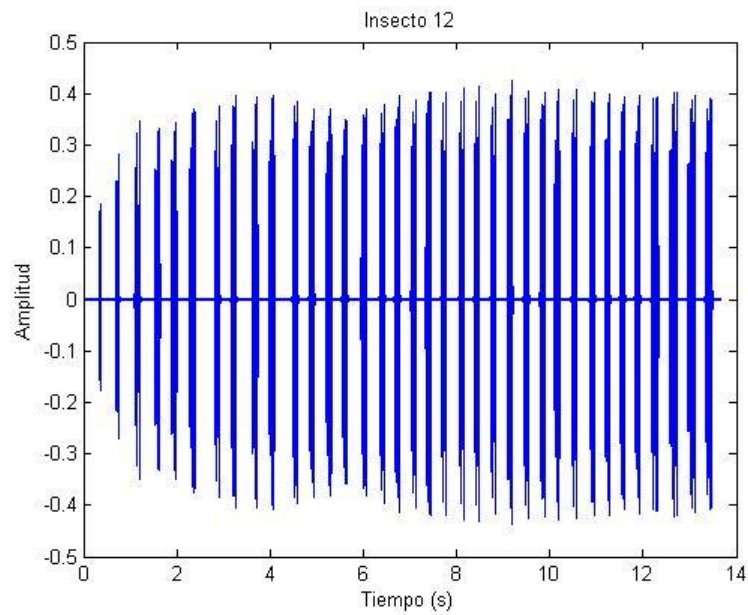


Figura 51 Insecto 12

Para este tipo de insecto tenemos un buen funcionamiento al usar valores de smoothness elevado como se muestra en la siguiente tabla con los fallos producidos:

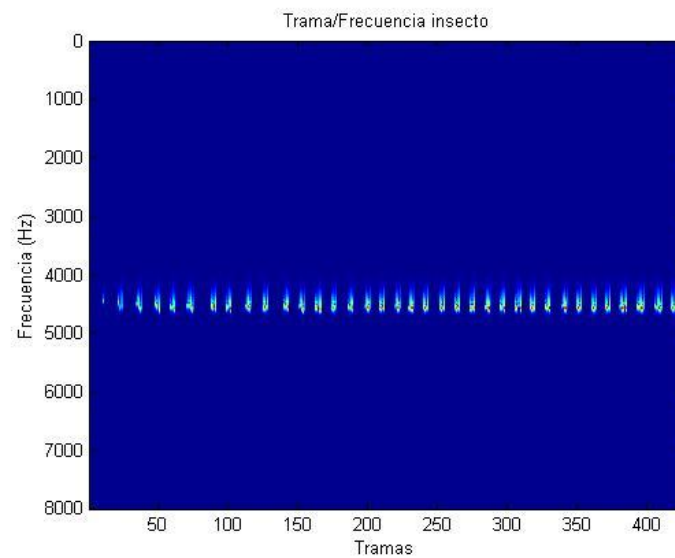


Figura 52 Espectrograma insecto 12

En la figura superior, vemos como se distribuye a lo largo del tiempo las frecuencias calculadas del insecto trama a trama.

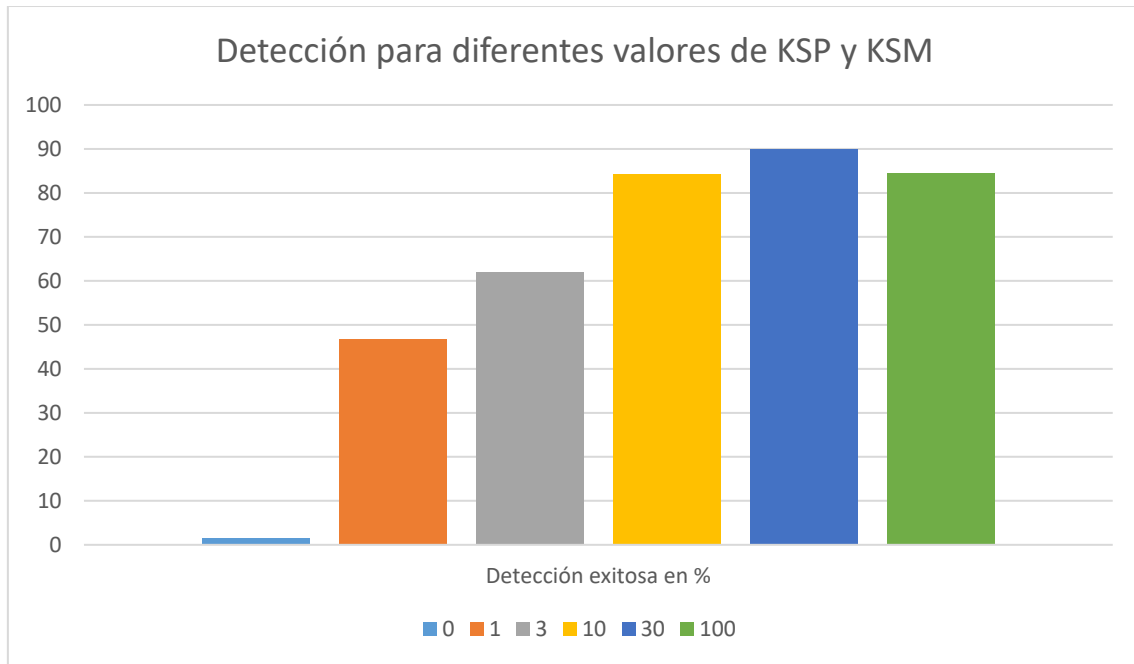


Figura 53 Detección exitosa para el insecto 12 con todos los tipos de ruido y para todos los valores de SNR. El número total de operaciones por cada valor de KSP y KSM es 336.

Los valores utilizados han sido 0 1 3 10 30 y 100, para smoothness y sparseness. En la gráfica que plasma un mejor funcionamiento para valores altos (10, 30 y 100) para smoothness y sparseness.

Por ejemplo, tenemos el insecto 12 con ruido de calle con un nivel de SNR de 0 dB:

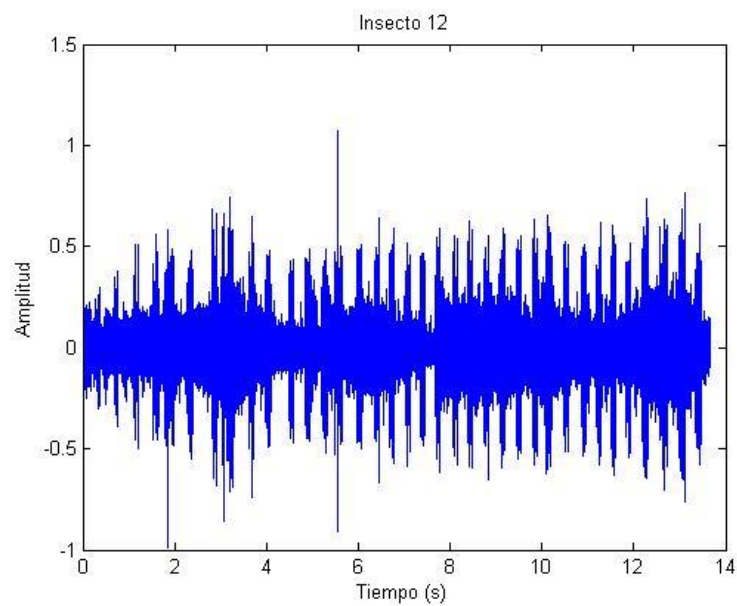


Figura 54 Insecto 12 con ruido de calle con un nivel de SNR de 0 dB

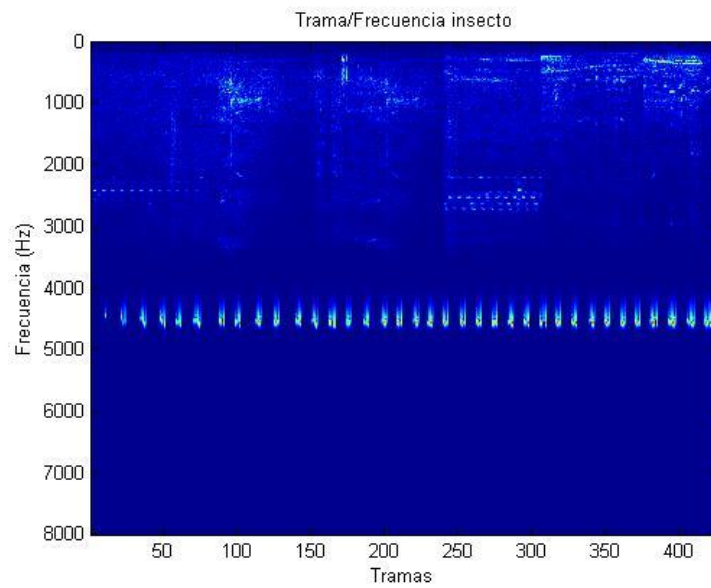


Figura 55 Espectrograma insecto 12 con ruido de calle con nivel de SNR de 0 dB

En el caso del insecto 14 tenemos la misma respuesta, donde sólo funciona el reconocimiento con valores de smoothness y sparseness alto (30 y 100).

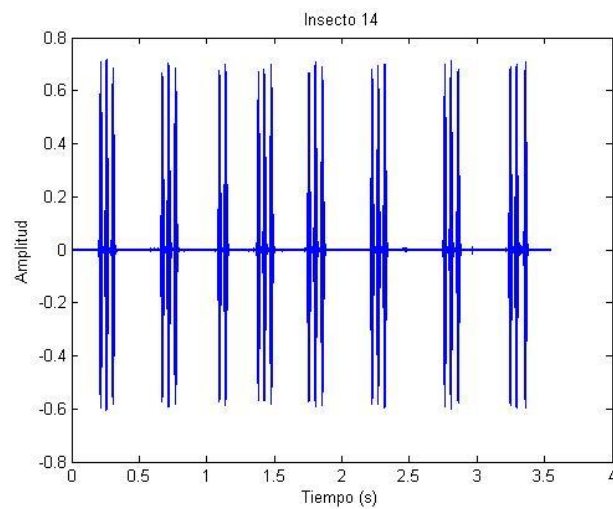


Figura 56 Insecto 14

La distribución de frecuencias a lo largo del fichero es la siguiente:

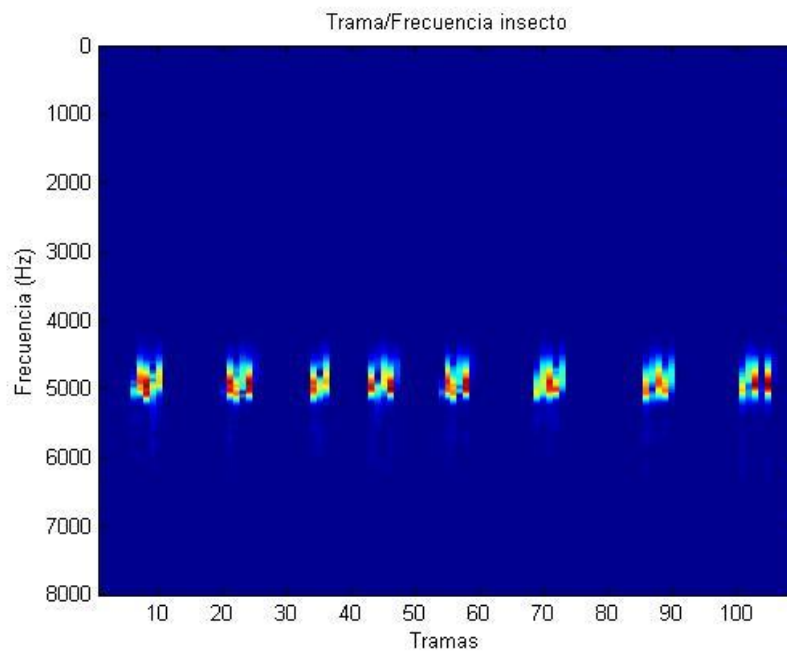


Figura 57 Espectrograma insecto 14

La tabla donde se muestran el porcentaje de éxito para cada valor de smoothness y sparseness con todos los tipos de ruido y para todos los niveles de SNR es el siguiente:

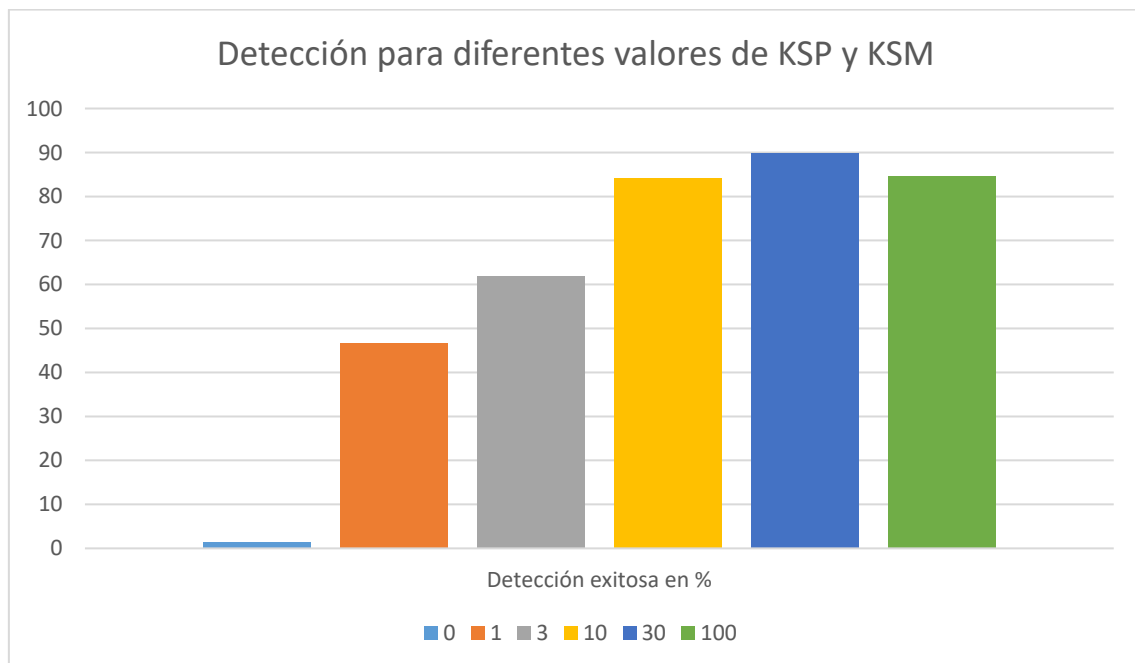


Figura 58 Porcentaje de éxito en detección del insecto 14 para todos los ruidos y todos los valores de SNR. El número total de operaciones por cada valor de KSP y KSM es 336.

El que sólo funcione con valores elevados de smoothness y sparseness indica la necesidad de suavidad en la activación de las bases para conseguir reconocer el insecto dado.

Por ejemplo, tenemos el insecto 14 con ruido de tren para un nivel de SNR de 0 dB:

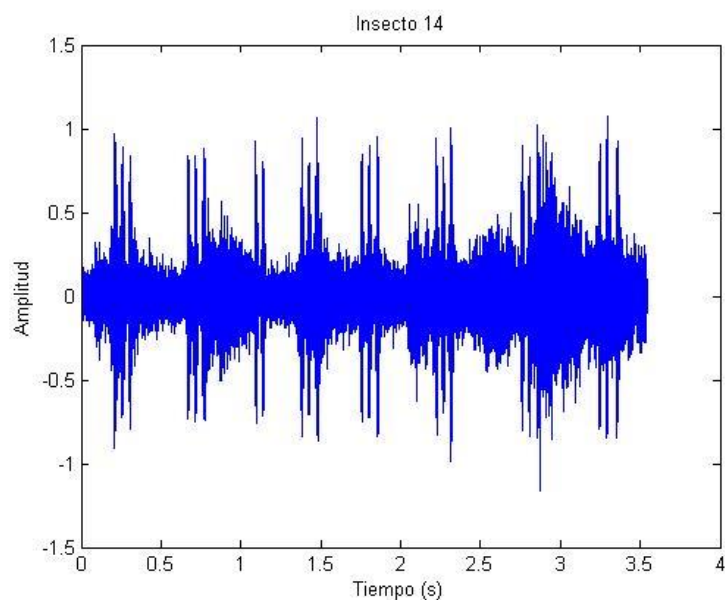


Figura 59 Insecto 14 con ruido de tren y un nivel de SNR de 0 dB

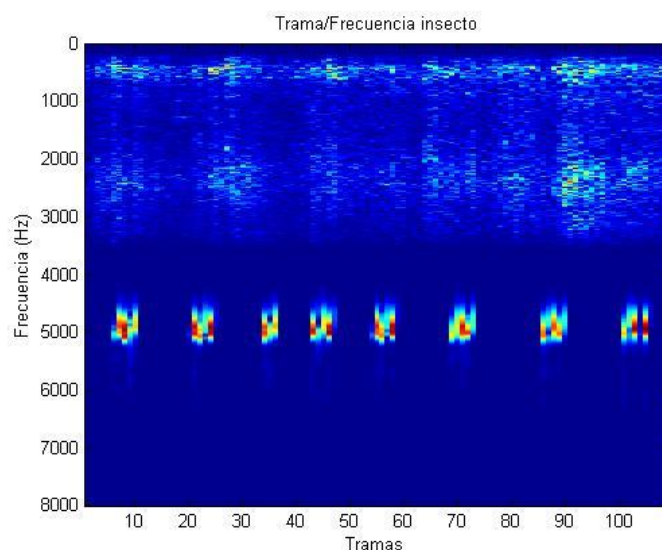


Figura 60 Espectrograma insecto 14 con ruido de tren 0 dB

Para el resto de insectos para todos los tipos de ruido, los valores más adecuados de smoothness y sparseness tienden a ser bajos. Para la mayoría el reconocimiento se produce con valores de 1 y 3.

Como hemos visto, con los tipos de ruido descritos anteriormente, el uso de 3 bases de ruido mejora el sistema entorno a un 3-4% para estos mismos casos.

3.3.3.1 Conclusiones añadiendo Noizeus con número fijo de bases y variación de valores de sparseness y smoothness

En esta prueba se ha detectado que el parámetro sparseness (escasez) no suele afectar al resultado de la identificación. Sin embargo el parámetro crítico es el de suavidad (smoothness) donde se tienen mejores resultados al utilizar un valor bajo en la gran mayoría de los casos, aunque ya hemos visto dos excepciones donde se mejora con valores altos.

Como conclusión, se determina como valor crítico del sistema de reconocimiento, el parámetro de smoothness (suavidad). Hemos visto que para valores 1 y 3 se produce una mayor identificación positiva comparado con el resto de valores y como consecuencia se procede a una nueva prueba con estos valores, donde se mantienen las bases de todos los tipos de ruido y las bases que caracterizan al insecto idénticas variándolas con el mismo valor y manteniendo fijo el valor de sparseness y smoothness.

3.3.4 Resultados añadiendo Noizeus manteniendo valores de sparseness y smoothness y modificando las bases de caracterización

En este apartado, se han realizado diferentes pruebas. En cada una de ellas, se ha mantenido el valor de smoothness y sparseness fijos variando las bases con las que se caracterizan el insecto y los ruidos. Las bases han sido las mismas para todos los elementos.

En la primera de ellas, se ha procedido a mantener los valores de smoothness y sparseness fijos a 1. Se han fijado unas bases para caracterizar los insectos iguales a las bases que se utilizarán para discriminar el ruido blanco, impulsivo y armónico. Se realiza un barrido para un rango de valores amplio de SNR para comprobar el comportamiento del sistema con las nuevas configuraciones. En concreto, se ha aplicado una SNR desde 20 dB a -20 dB con saltos de 5 dB para cada tipo de ruido con los diferentes valores de bases que se han utilizado (3, 5 y 7).

A continuación, se muestra una tabla con el porcentaje de éxito en la identificación de insectos:

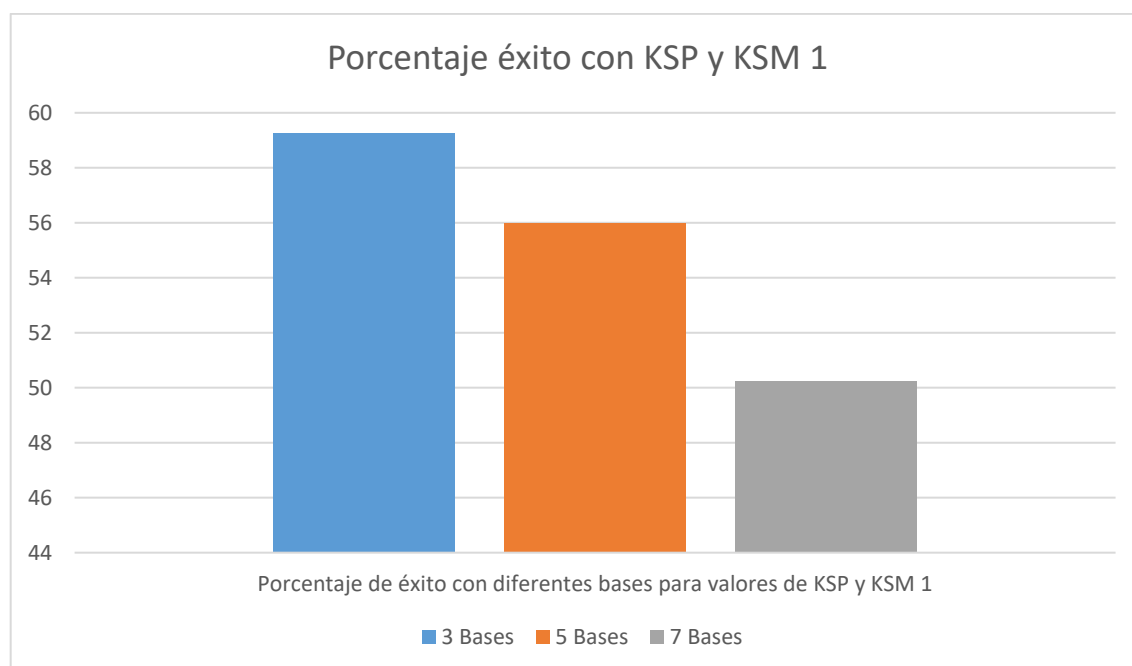


Figura 61 Porcentaje de éxito en la identificación de insectos para diferentes números de bases con KSP y KSM fijo a 1. Se toman en cuenta todos los tipos de ruido y los valores de SNR (20 hasta -20dB). El número total de operaciones por cada valor de bases es 1656.

En el gráfico, vemos como el conjunto de 3 bases para todos los tipos de ruido como para las bases de caracterización del insecto se obtienen mejores resultados, consiguiendo detectar el insecto con hasta un nivel de SNR de -20 dB en algunos casos. Como norma general, el sistema funciona correctamente hasta una media de -5dB para esta configuración. Cierto es que dependiendo del insecto y del tipo de ruido empleado este nivel se vea disminuido o incluso aumentado.

En el caso del uso de 5 y 7 bases, el porcentaje de éxito se sitúa por encima del 50%, siendo más factible utilizar 3 bases. Como muestra la figura, la diferencia entre el uso de 3 bases y 5 bases es de un 4% aproximadamente.

En el caso del uso de 7 bases, tenemos un porcentaje de éxito del 50% para el número total de operaciones realizadas.

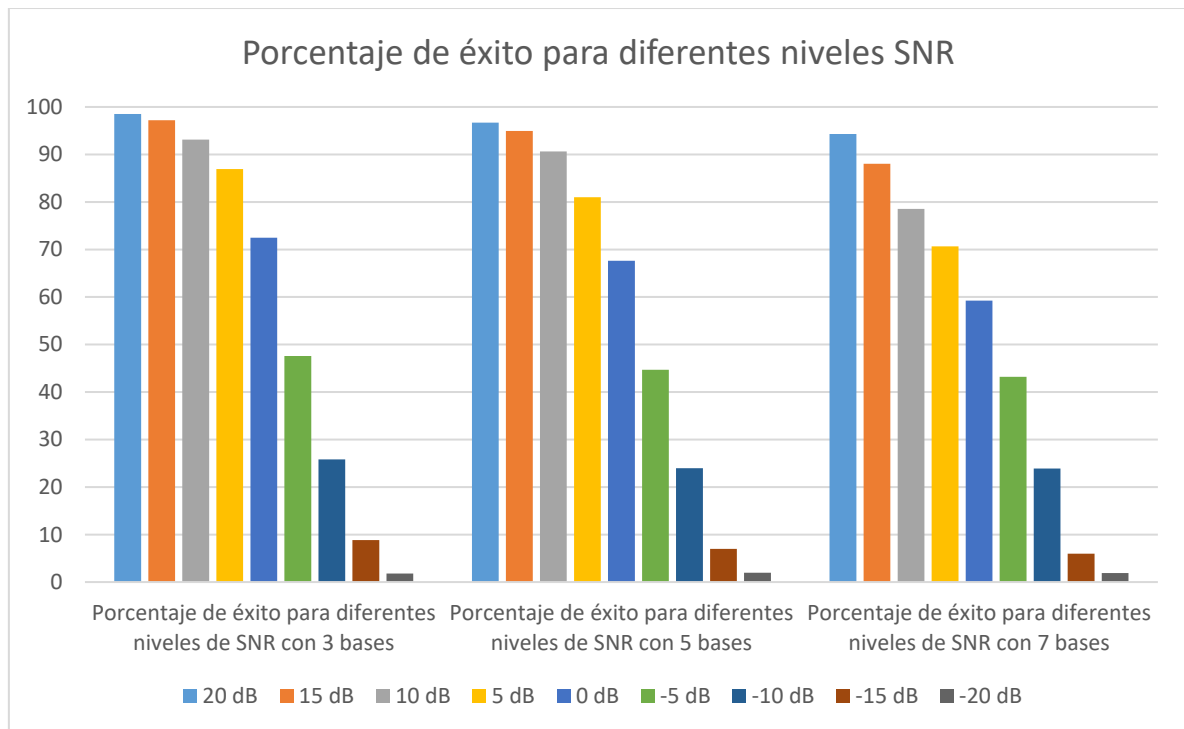


Figura 62 Porcentaje de éxito en la identificación de insectos para diferentes niveles de SNR diferentes números de bases con KSP y KSM fijo a 1. El número total de operaciones por cada valor de SNR es 368.

En la segunda fase, se ha realizado el mismo procedimiento anterior variando el valor de smoothness y sparseness a 3 para comprobar cuál de los dos valores funciona mejor. Se han aplicado para el resto de elementos los mismos valores, es decir para la SNR se varía desde 20 dB hasta -20 dB con saltos de 5 dB y las bases de los elementos de ruido y caracterización del insecto se varían con los siguientes: 3, 5 y 7, dejando siempre los valores iguales en todas las bases.

Como se ve en la ilustración inferior, se vuelven a repetir los mismos resultados anteriores, donde con 3 bases se produce un mejor reconocimiento en comparación con el uso de las demás bases. Así mismo, se repite el funcionamiento con las demás bases, 5 y 7 para valores de SNR elevados donde no existe tanta dificultad a la hora de diferenciar el insecto con el ruido de fondo.

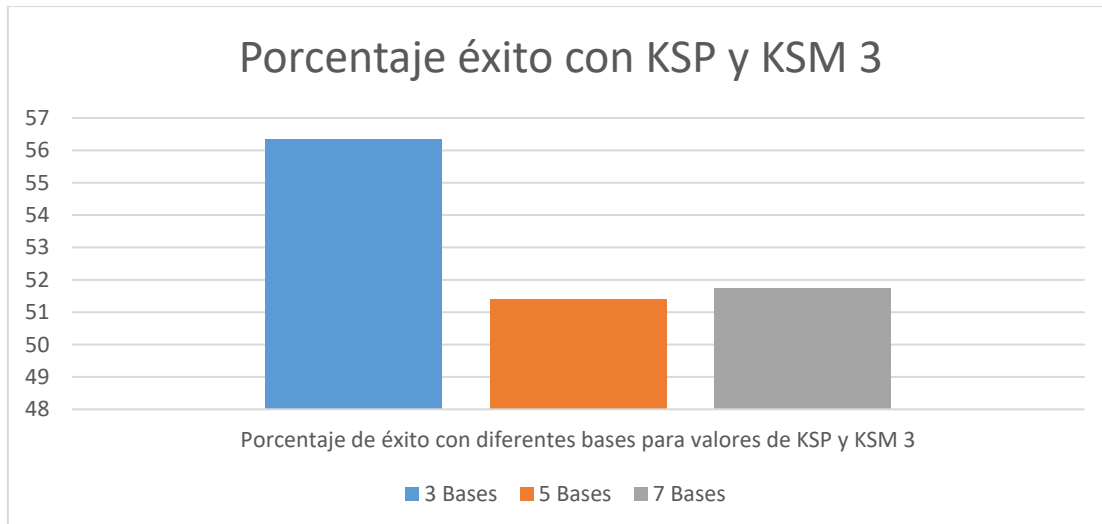


Figura 63 Porcentaje de éxito en la identificación de insectos para diferentes números de bases con KSP y KSM fijo a 3. Se toman en cuenta todos los tipos de ruido y los valores de SNR (20 hasta -10dB). El número total de operaciones por cada valor de bases es 1656.

De la misma manera, se han realizado las pruebas pertinentes para los valores 10, 30 y 100 para smoothness y sparseness. En las siguientes gráficas se muestran los resultados obtenidos para las diferentes bases utilizadas. Vamos a ver como para valores elevados, se tiene un mejor comportamiento del sistema con una configuración de 5 bases aunque la tasa de éxito disminuye considerablemente.

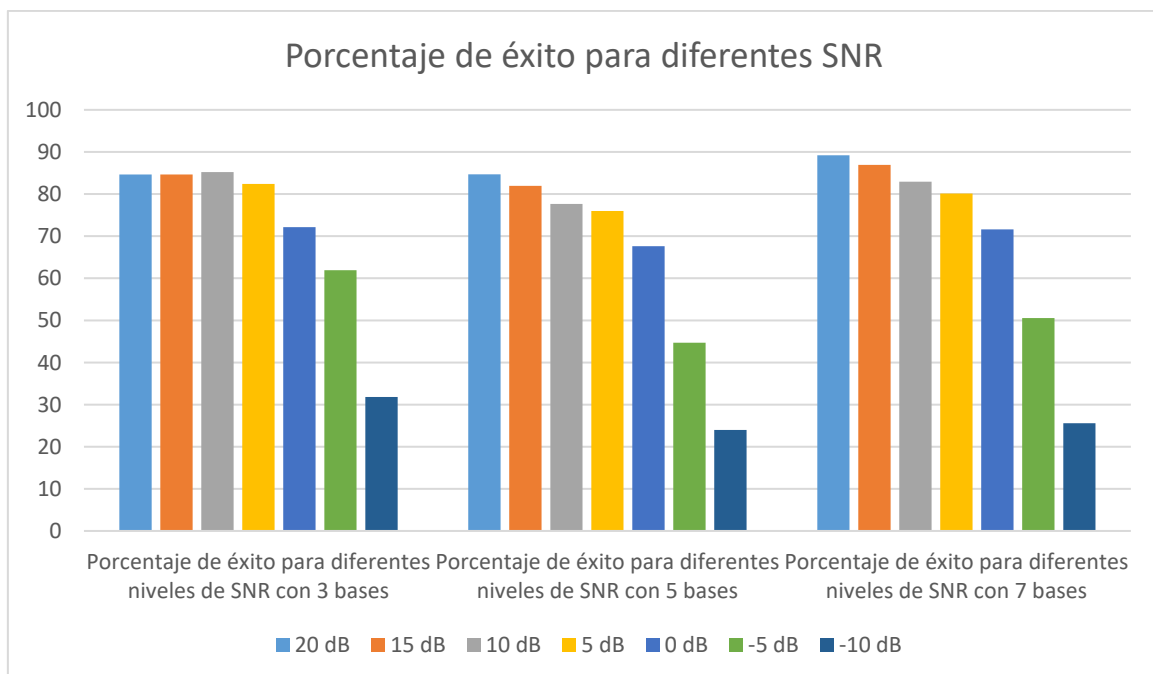


Figura 64 Porcentaje de éxito en la identificación de insectos para diferentes niveles de SNR diferentes números de bases con KSP y KSM fijo a 3. El número total de operaciones por cada valor de SNR es 368.

Para el caso de utilizar un valor de smoothness y sparseness igual a 10 tenemos la siguiente gráfica:

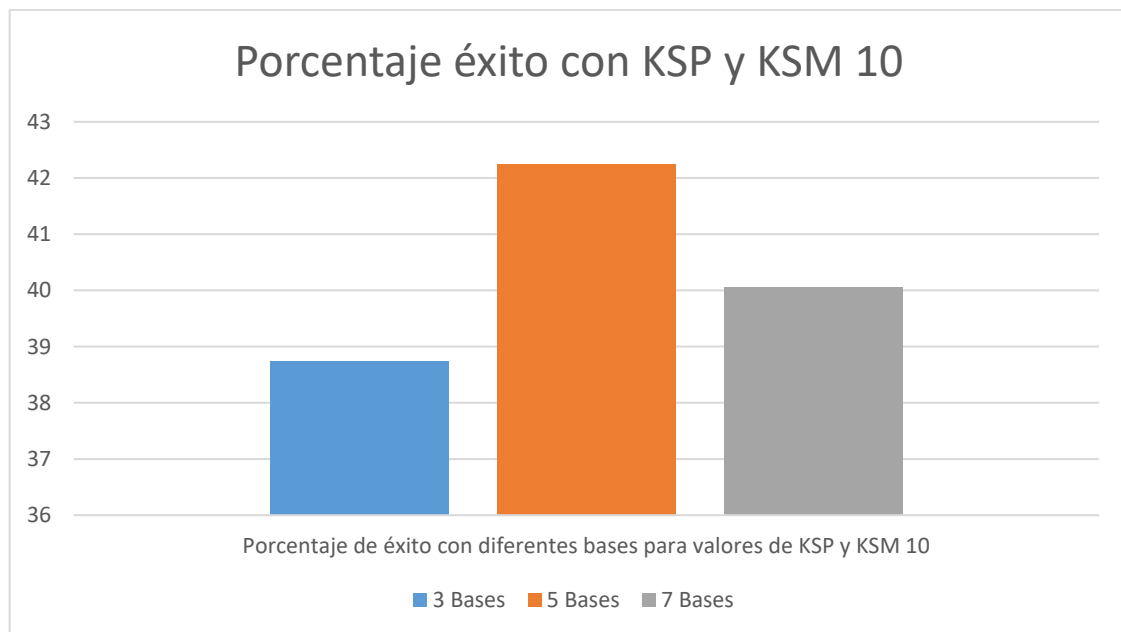


Figura 65 Porcentaje de éxito en la identificación de insectos para diferentes números de bases con KSP y KSM fijo a 10. Se toman en cuenta todos los tipos de ruido y los valores de SNR (20 hasta -10dB). El número total de operaciones por cada valor de bases es 1288.

Podemos ver como el uso de 5 bases para valores elevados de sparseness y smoothness, en comparación con el uso de otro número de bases, mejora el comportamiento, aunque la tasa de éxito se encuentra por debajo del 50 %.

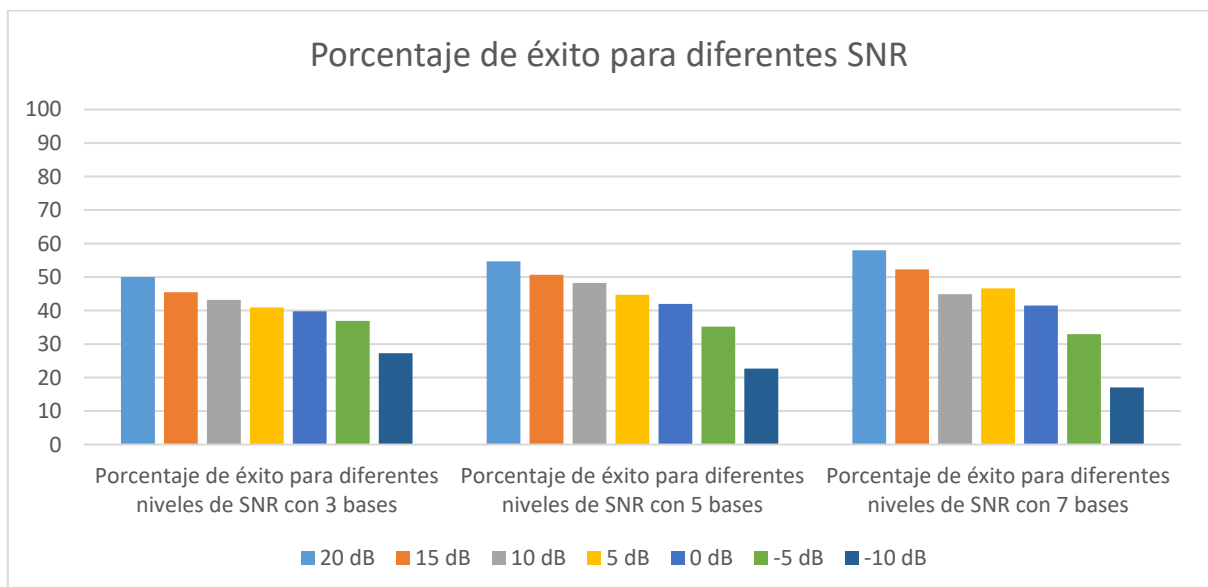


Figura 66 Porcentaje de éxito en la identificación de insectos para diferentes niveles de SNR diferentes números de bases con KSP y KSM fijo a 10. El número total de operaciones por cada valor de SNR es 368.

A continuación se muestran los resultados para las diferentes bases con un valor para smoothness y sparseness igual a 30.

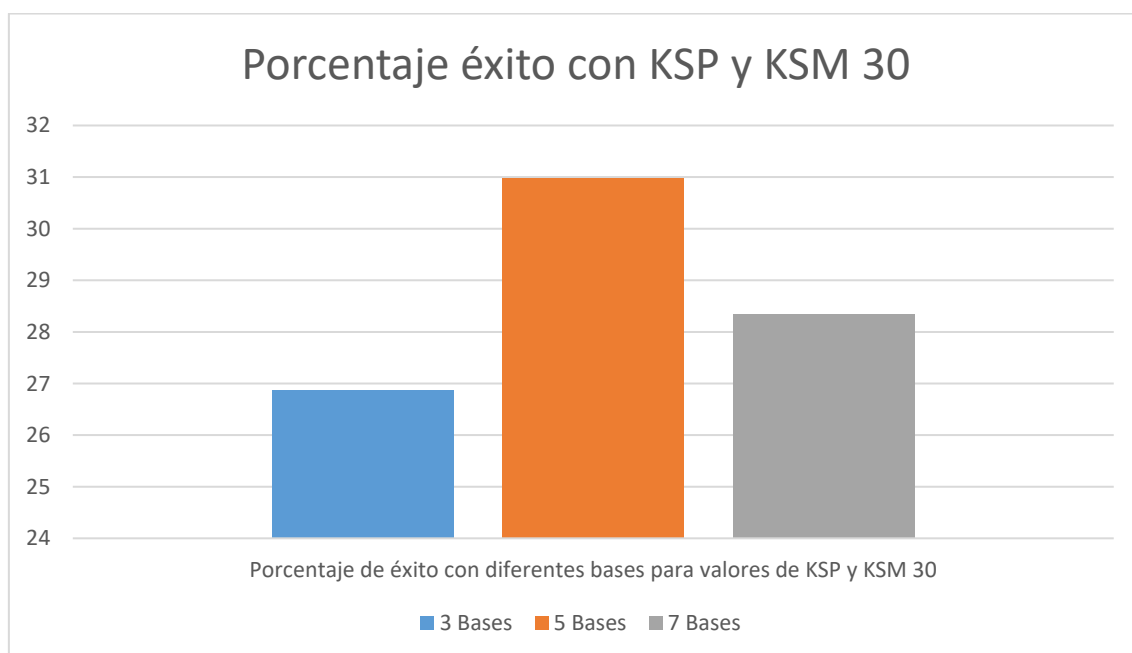


Figura 67 Porcentaje de éxito en la identificación de insectos para diferentes números de bases con KSP y KSM fijo a 30. Se toman en cuenta todos los tipos de ruido y los valores de SNR (20 hasta -10dB). El número total de operaciones por cada valor de bases es 1288.

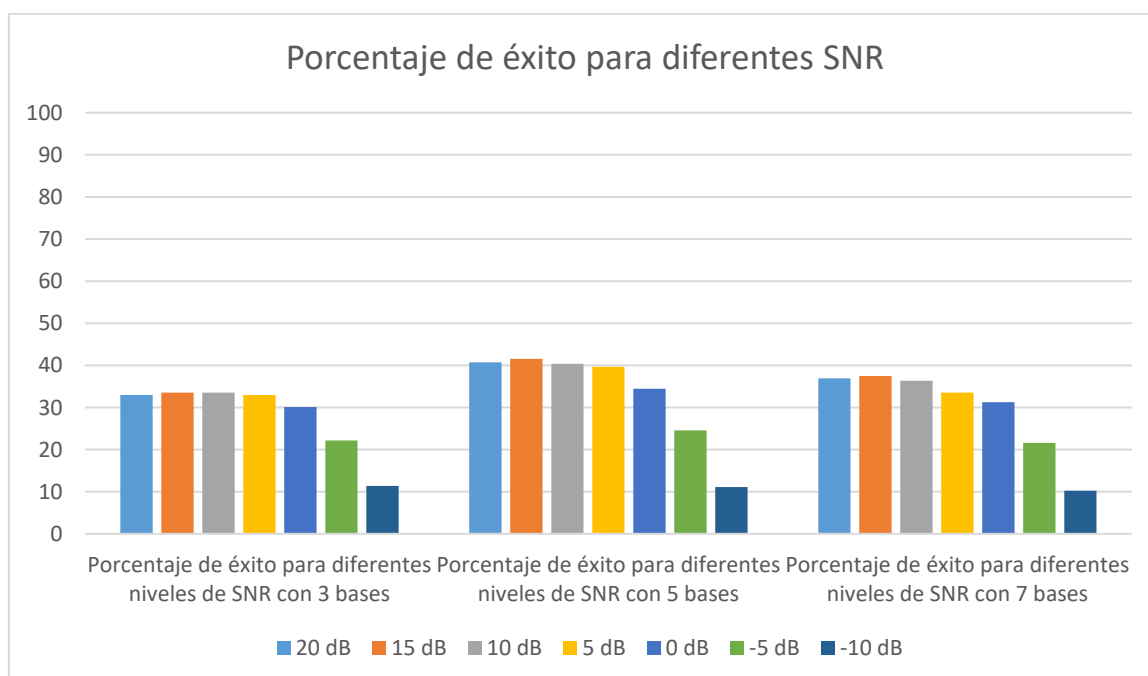


Figura 68 Porcentaje de éxito en la identificación de insectos para diferentes niveles de SNR diferentes números de bases con KSP y KSM fijo a 30. El número total de operaciones por cada valor de SNR es 368.

Por último, tenemos la gráfica usando un valor igual a 100 para smoothness y sparseness donde de nuevo vuelve a funcionar mejor en el sistema una configuración de 3 bases de caracterización, con tasas de éxito bastante bajas:

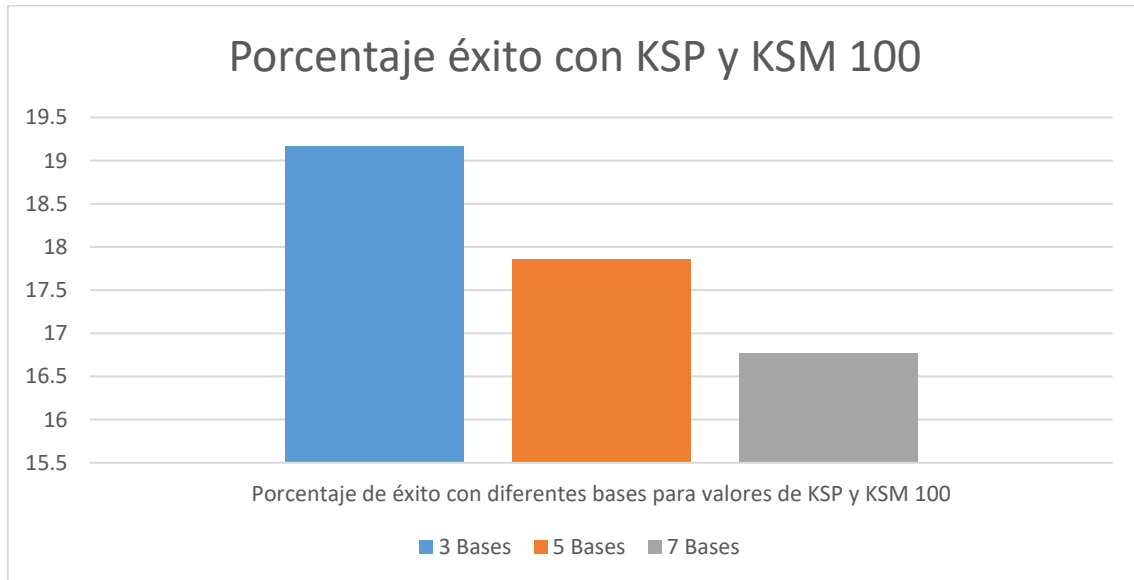


Figura 69 Porcentaje de éxito en la identificación de insectos para diferentes números de bases con KSP y KSM fijo a 100. Se toman en cuenta todos los tipos de ruido y los valores de SNR (20 hasta -10dB). El número total de operaciones por cada valor de bases es 1288.

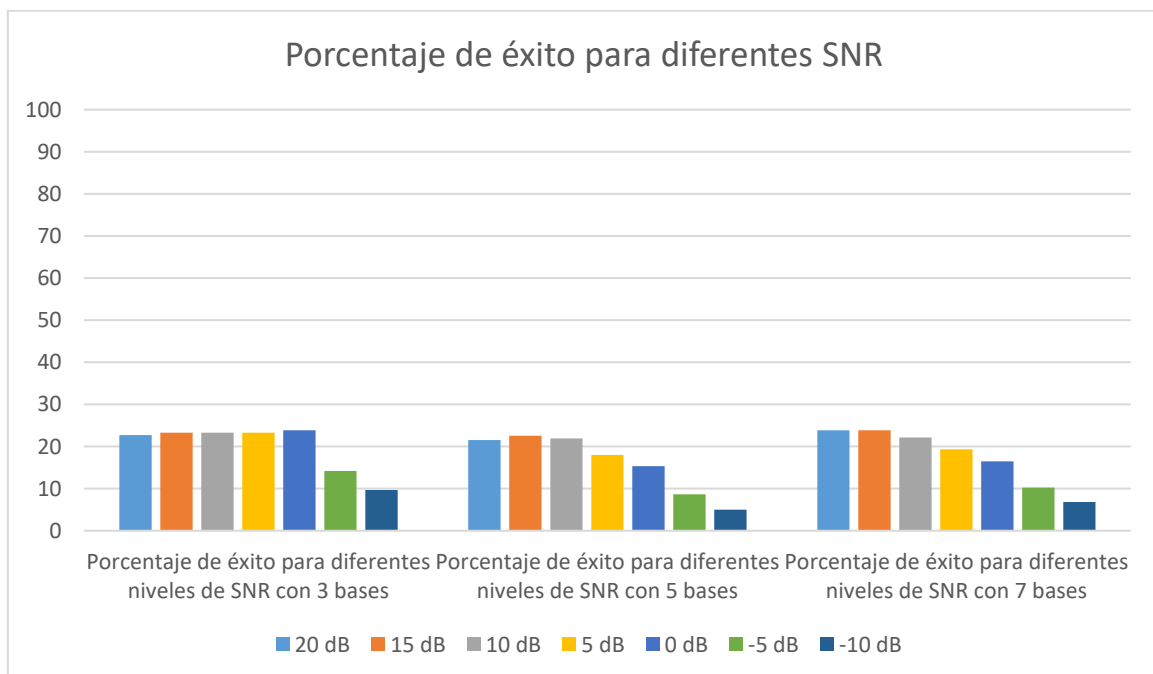


Figura 70 Porcentaje de éxito en la identificación de insectos para diferentes niveles de SNR diferentes números de bases con KSP y KSM fijo a 100. El número total de operaciones por cada valor de SNR es 368.

Podemos afirmar entonces, que para un número de bases de caracterización igual a 3 junto con valores bajos de smoothness y sparseness (1 y 3), se tienen mejores resultados a la hora de identificar el insecto en cuestión.

Una vez detectado para que valores se comporta mejor las diferentes bases y sacado en claro que el mejor valor ha sido 1 para smoothness y sparseness, se ha realizado una prueba para comprobar la viabilidad al utilizar bases diferentes a las estudiadas anteriormente. Se ha realizado una prueba con los valores de sparseness y smoothness fijos a 1 para los valores de bases de caracterización siguientes: 1,2 y 4. A continuación se muestran los resultados obtenidos junto al estudio con los datos anteriores mostrados.

Comprobamos como para 1, 3 y 5 bases se obtienen mejores resultados. Por lo que nos centraremos en ellos realizando un análisis de su comportamiento para cada nivel de SNR. La diferencia de nivel entre 1 y 3 es de un 1% siendo una diferencia no muy elevada y teniendo prácticamente los mismos resultados.

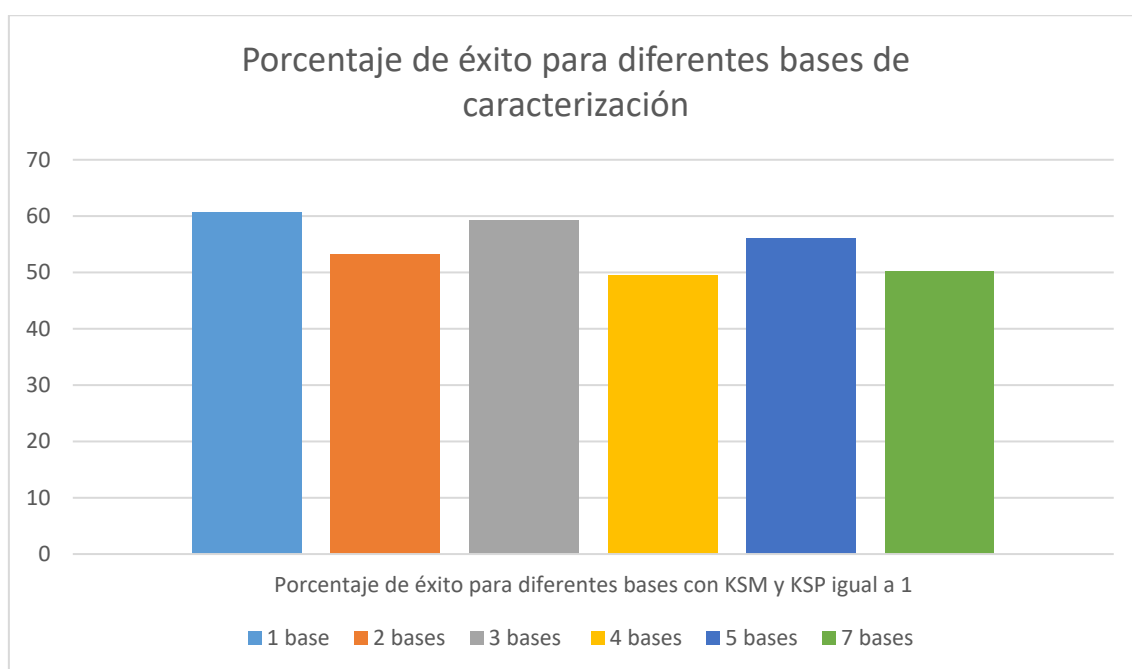


Figura 71 Porcentaje de éxito en la identificación de insectos para diferentes números de bases con KSP y KSM fijo a 1. Se toman en cuenta todos los tipos de ruido y los valores de SNR (20 hasta -20dB). El número total de operaciones por cada valor de bases es 1656.

3.3.4.1 Conclusiones añadiendo Noizeus con valores fijos de sparseness y smoothness y variación de número de bases de caracterización

En este apartado, se han obtenido buenos resultados con la utilización de 1 y 3 bases de caracterización de insectos y ruido para unos valores determinados de sparseness y smoothness.

Se ha visto que un mayor número de bases no significa que haya una mejora en el reconocimiento, todo lo contrario, se produce un descenso notable de la viabilidad del sistema. Es cierto que para valores muy bajos de SNR, en determinadas ocasiones, el sistema responde de manera más eficiente con mayor número de bases, en nuestro caso 5, pero en demás situaciones y como norma general con 1 o 3 bases es suficiente y la mejor de las opciones.

Por último, entre los valores de smoothness y sparseness utilizados, se encuentran dos valores donde el reconocimiento se hace más notable. En concreto hablamos de los valores 1 y 3. Como hemos visto en los gráficos, con la utilización del valor 1 para smoothness y sparseness, se obtienen mejores resultados en comparación con el valor 3.

3.3.5 Estudio aislado de insectos 13 y 17

Como ya se ha mencionado, existen ciertos insectos donde se produce un error en el reconocimiento para casi todos los valores de combinación usados. Estos insectos, el 13, el 17 y el 22, al aplicar el software de reconocimiento y tener como configuración alguna base de ruido blanco no se consigue un resultado correcto.

A continuación se muestra la forma de onda de los insectos 13, 17 y 22, así como su espectrograma:

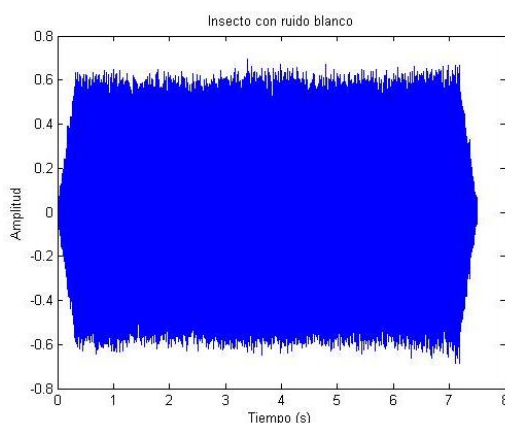


Figura 72 Insecto 13

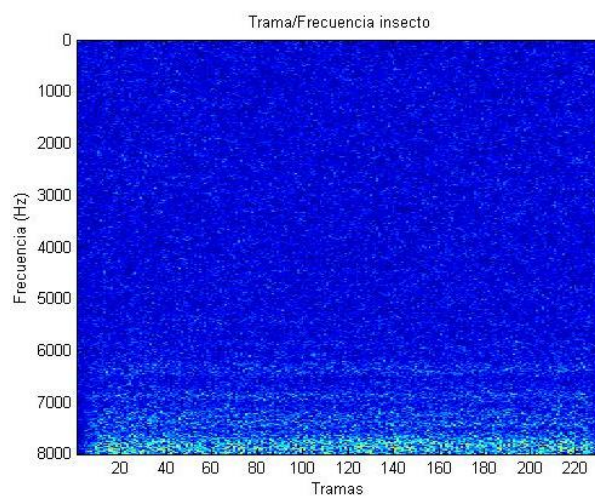


Figura 73 Espectrograma insecto 13

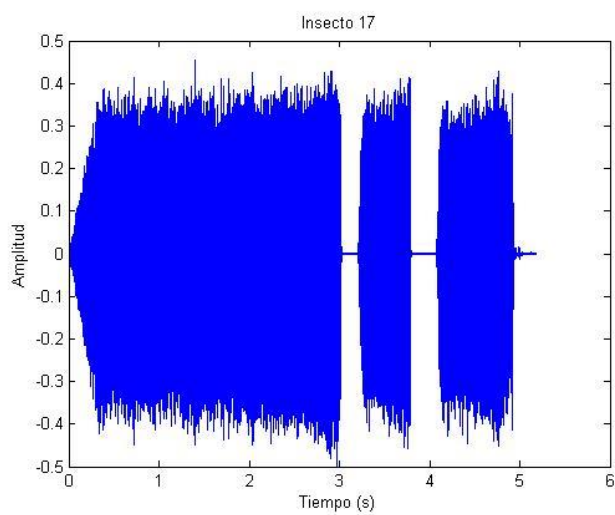


Figura 74 Insecto 17

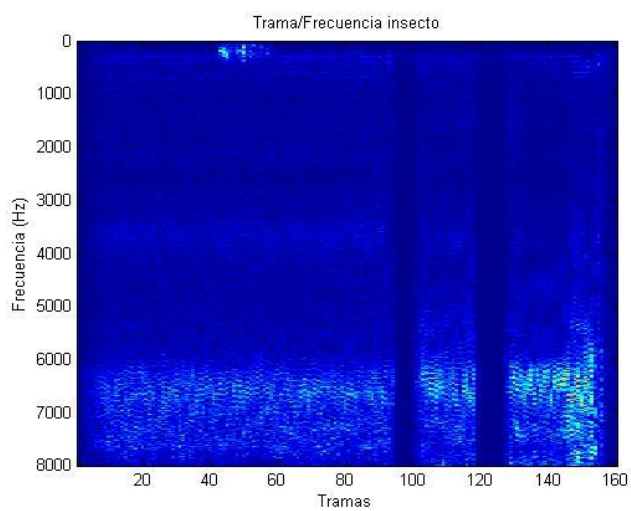


Figura 75 Espectrograma insecto 17

Podemos apreciar una distribución de frecuencias de ambos insectos les muy similar.

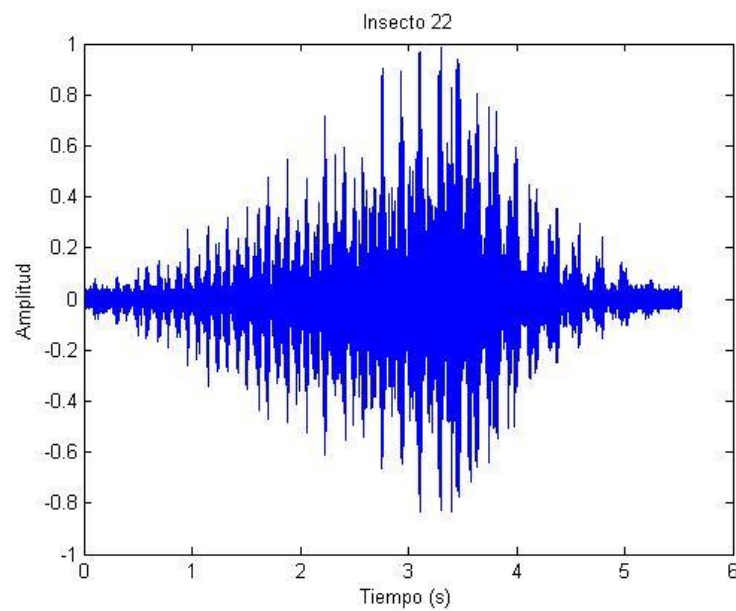


Figura 76 Insecto 22

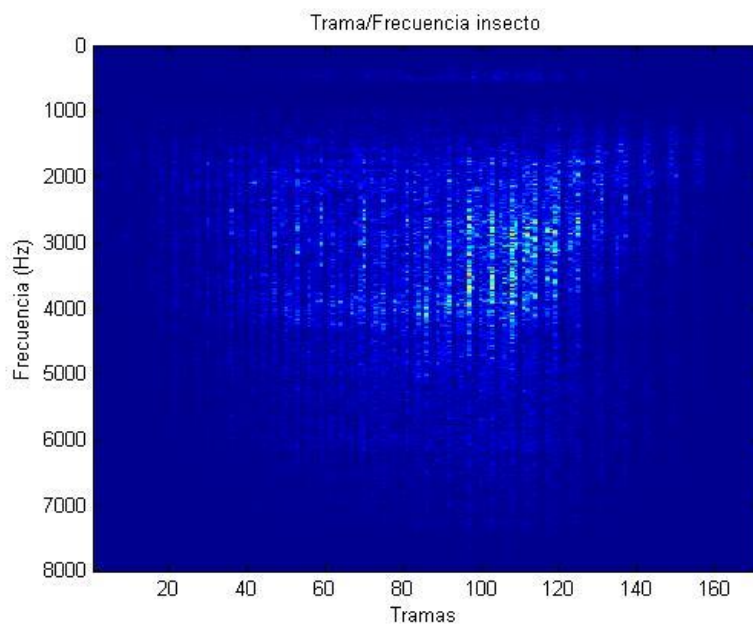


Figura 77 Espectrograma insecto 22

Como ejemplo, tenemos el insecto 17 junto con ruido de conversación (babble) con un nivel de SNR de 0 dB, no se logra detectar para ningún valor de sparseness y smoothness con un número de bases de caracterización igual a 5. El sonido generado se muestra en la imagen inferior.

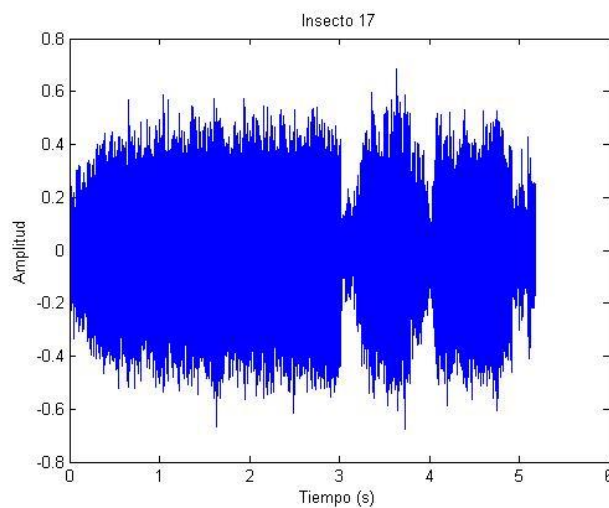


Figura 78 Insecto 17 con ruido de conversación y un nivel de SNR de 0 dB

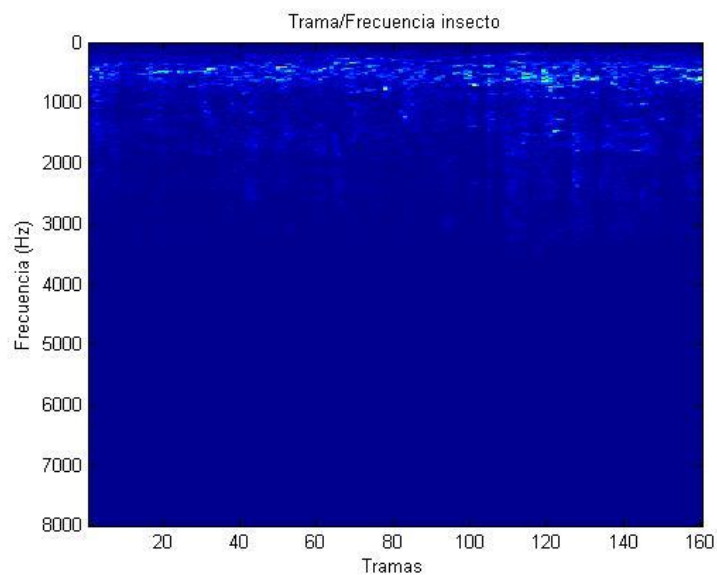


Figura 79 Espectrograma insecto 17 con ruido de conversación y un nivel de SNR de 0 dB

Vemos como prácticamente ha desaparecido la distribución de frecuencias original del insecto y no se es capaz de llevar a cabo la detección.

Otro ejemplo, en este caso del insecto 13, tenemos añadido ruido de calle con un nivel de SNR de 10 dB:

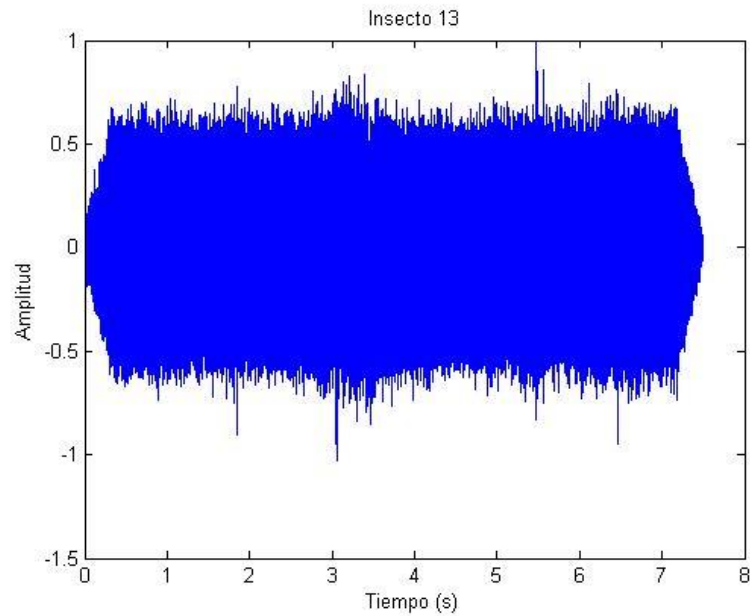


Figura 80 Insecto 13 con ruido de calle con un nivel de SNR de 10 dB

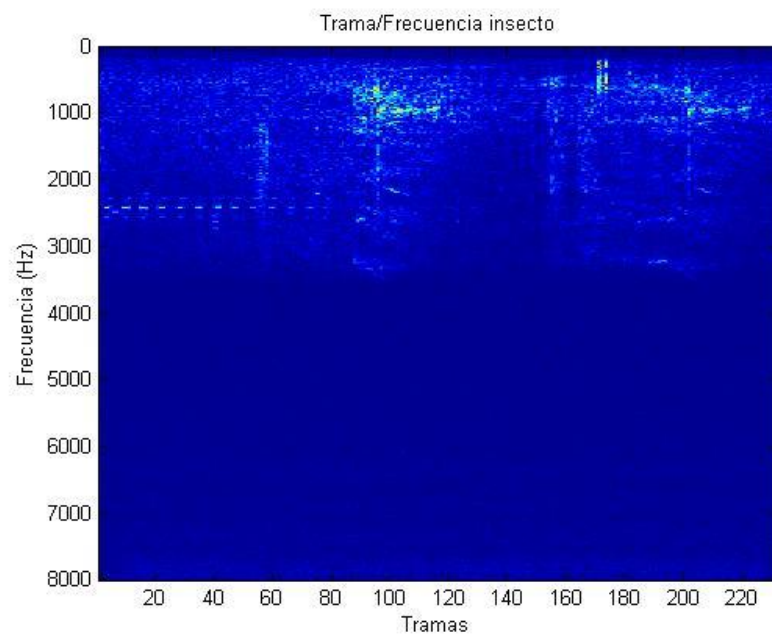


Figura 81 Espectrograma insecto 13 con ruido de calle para un nivel de SNR de 10 dB

De nuevo, se observa como el espectro de nuestro insecto original se modifica casi completamente al añadirle un tipo de ruido y conlleva la identificación negativa de éste.

Se procede a estudiar los resultados de los insectos 13, 17 y 22 de manera aislada para comprobar que tipo de combinación es la más correcta para utilizar en su caso.

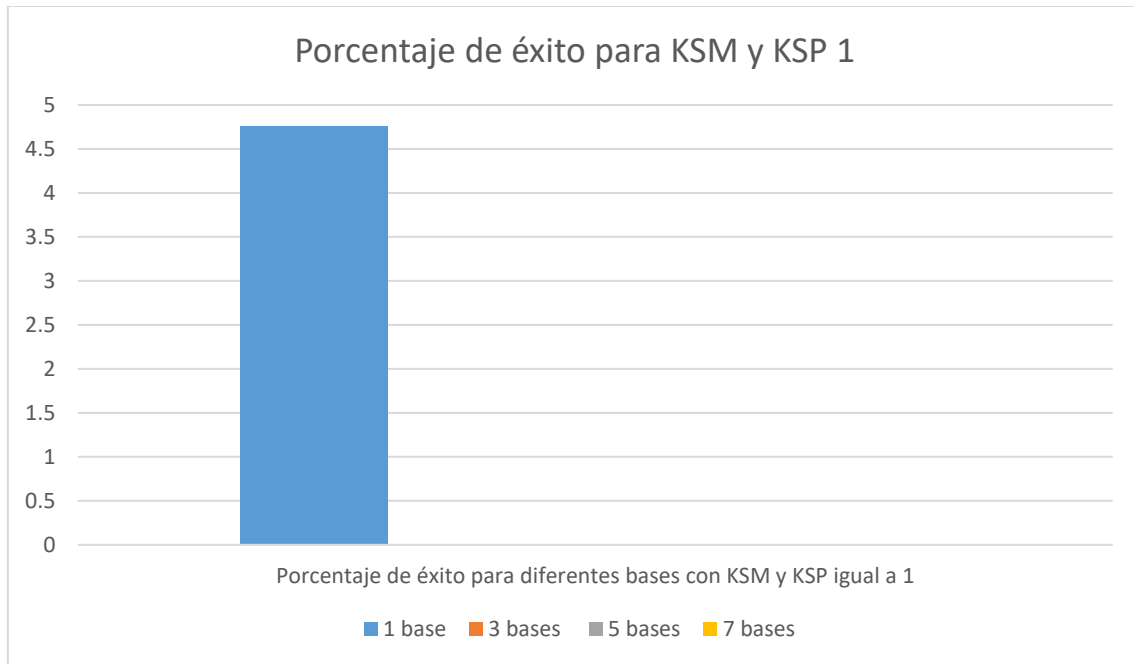


Figura 82 Gráfico de porcentaje de éxito en el reconocimiento de insectos con una configuración para los insectos y los tipos de ruido y un valor 1 para sparseness y smoothness. El total de operaciones por cada nivel de SNR han sido 168.

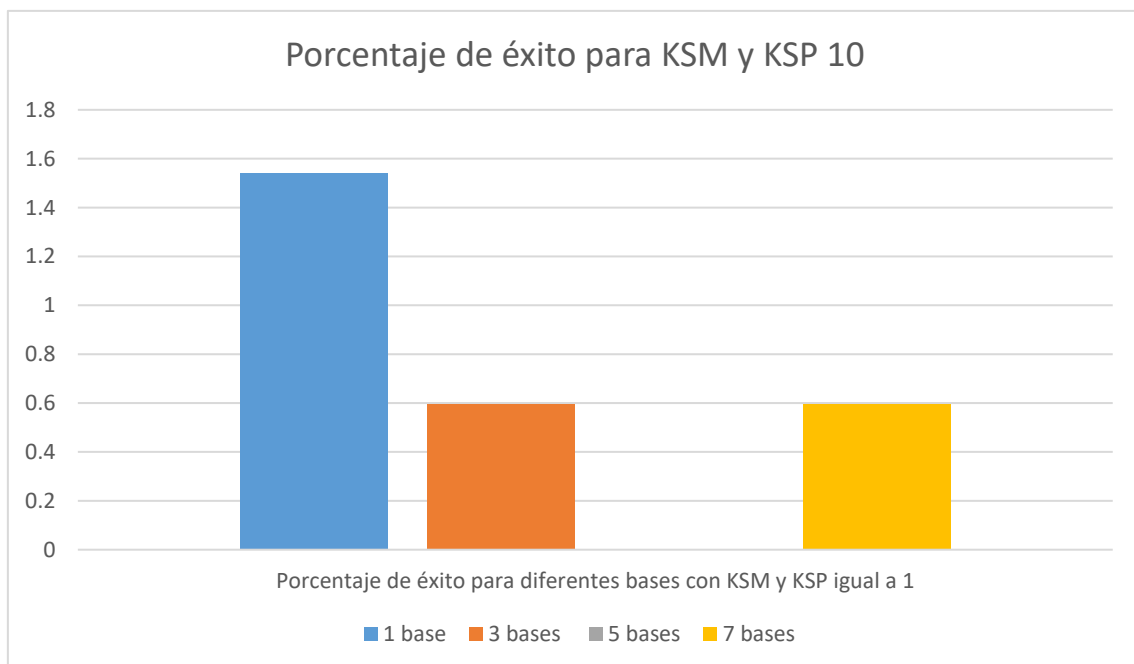


Figura 83 Gráfico de porcentaje de éxito en el reconocimiento de insectos con una configuración para los insectos y los tipos de ruido y un valor 10 para sparseness y smoothness. El total de operaciones por cada nivel de SNR han sido 168.

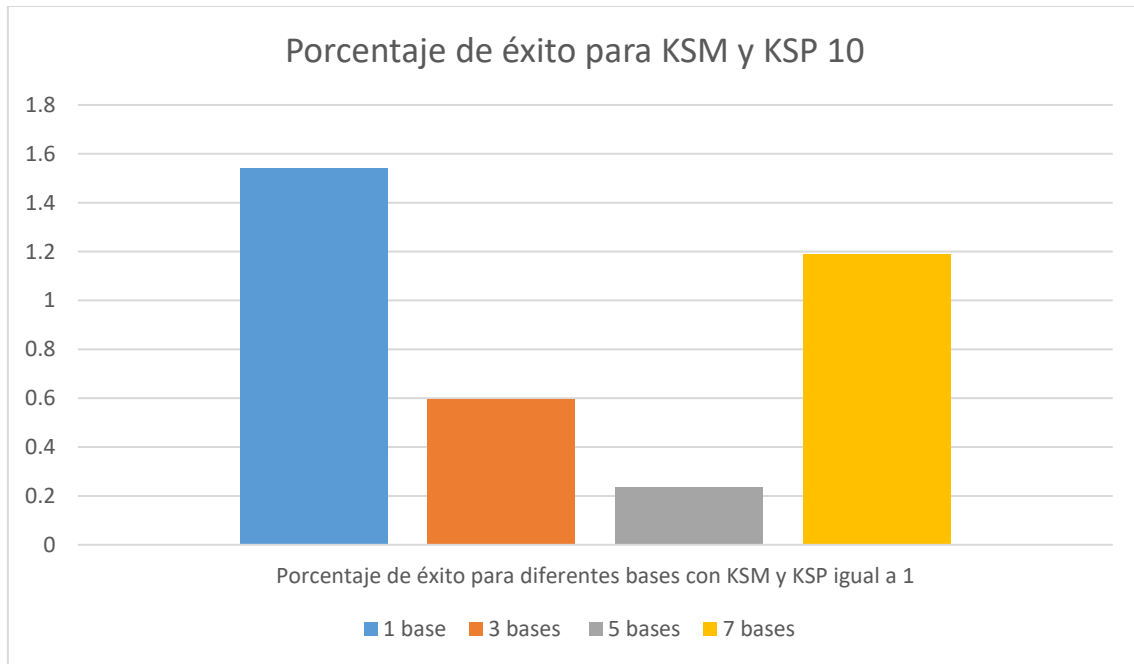


Figura 84 Gráfico de porcentaje de éxito en el reconocimiento de insectos con una configuración para los insectos y los tipos de ruido y un valor 100 para sparseness y smoothness. El total de operaciones por cada nivel de SNR han sido 168.

Como vemos, la mejor configuración para estos insectos es utilizar una base de caracterización y un valor 1 para smoothness y sparseness, aunque la detección sea del 4%

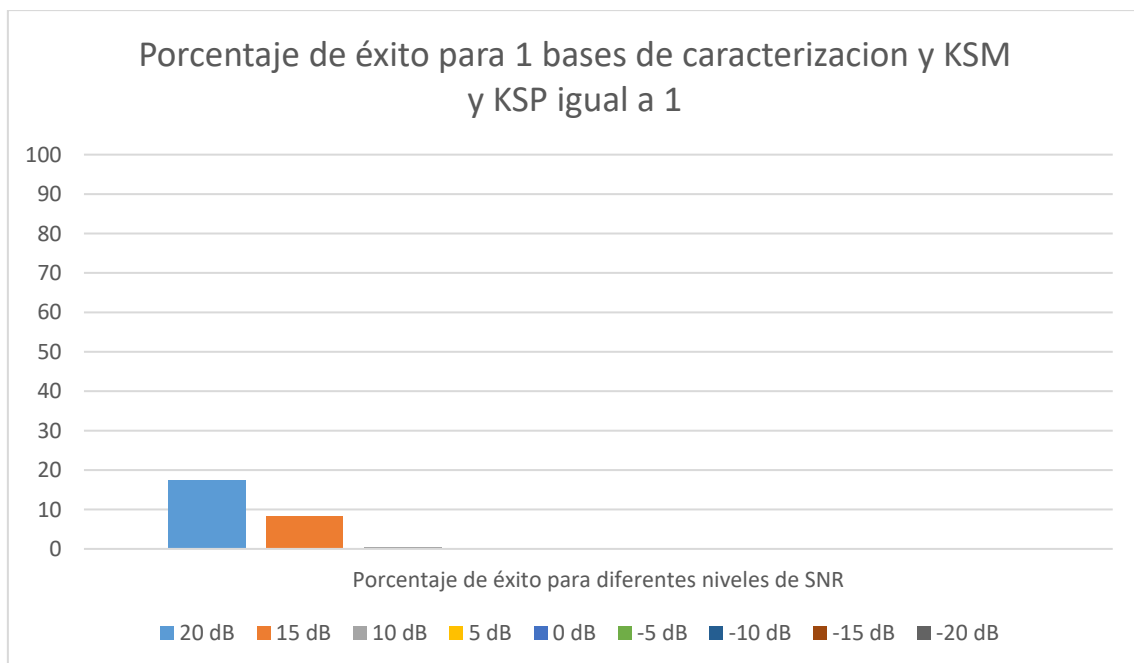


Figura 85 Gráfico de porcentaje de éxito en el reconocimiento de insectos con una configuración de 1 base de caracterización para los insectos y los tipos de ruido y un valor 1 para sparseness y smoothness. El total de operaciones por cada nivel de SNR han sido 264.

Podemos ver en la figura superior, que el reconocimiento sólo se produce en niveles de muy bajo ruido donde la SNR es alta (20 y 15 dB). Como vemos el éxito es muy escaso para estos niveles.

3.3.5.1 Conclusiones del estudio de insectos 13, 17 y 22

En relación a los insectos 13, 17 y 22, se ha visto que no se detecta para casi ninguna configuración dada y se demuestra en sus espectrogramas la modificación de ellos al añadir algún tipo de ruido perdiendo sus características iniciales.

Sólo es posible detectarlos con niveles muy altos de SNR y en muy pocas ocasiones. Esto es debido, como podemos ver en sus espectrogramas, que la potencia es mayor que en otros casos debido al canto continuo del insecto. Al añadirle ruido, se le está añadiendo un valor mucho más elevado que las ocasiones anteriores produciendo una degradación de la señal más brusca e imposibilitando la identificación de estos.

CONCLUSIONES

Como conclusiones del estudio de identificación de insectos mediante técnicas NMF, se puede afirmar que es un sistema muy robusto logrando identificar a los insectos con niveles bajos de SNR. Se tiene un buen porcentaje de éxito en la detección para una combinación de 3 bases de caracterización para los diferentes tipos de ruido e insecto y un valor de sparseness y smoothness igual a 1.

En la siguiente figura, se muestran el porcentaje de éxito en el reconocimiento de insectos para la configuración general, en el que se incluyen todos los insectos (a excepción del insecto 13, 17 y 22) y los distintos tipos de ruido. Se muestran por niveles de SNR:

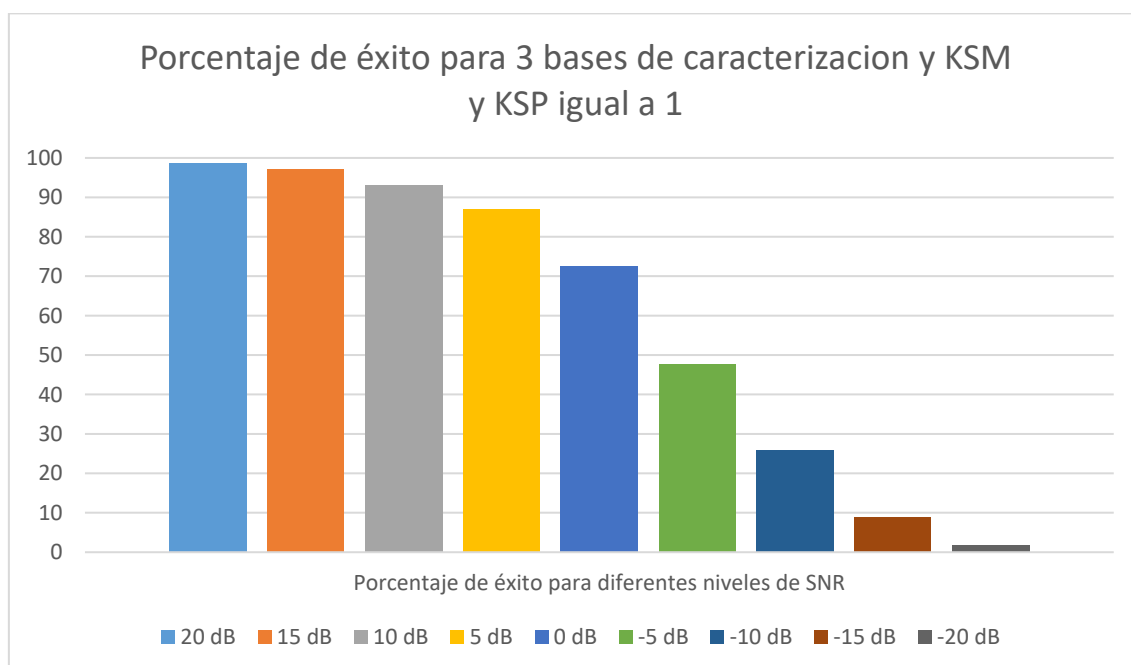


Figura 86 Gráfico de porcentaje de éxito en el reconocimiento de insectos con una configuración de 3 bases de caracterización para los insectos y los tipos de ruido y un valor 1 para sparseness y smoothness. El total de operaciones por cada nivel de SNR han sido 2024.

Podemos observar un muy buen comportamiento hasta 0 dB. Una vez tenemos un nivel de SNR negativo se degrada el éxito obtenido, teniendo para -5 dB un 47% de éxito, debido a que tenemos ya un importante nivel de ruido superior al sonido del insecto.

Si comparamos la opción del uso de 3 bases de caracterización y el uso de 5 para mismos valores de sparseness y smoothness, vemos como el comportamiento es muy similar aunque tenemos una diferencia aproximada del 3% entre ambos resultados para niveles de SNR elevados (20, 15 y 10 dB) y para el resto de niveles de SNR positivos, se tiene un porcentaje mayor (5-8%). Sin embargo, al tener niveles de SNR negativos vemos

como la diferencia se reduce teniendo un 2-3% de diferencia entre ambos resultados, dando a entender un comportamiento similar para valores de SNR bajos.

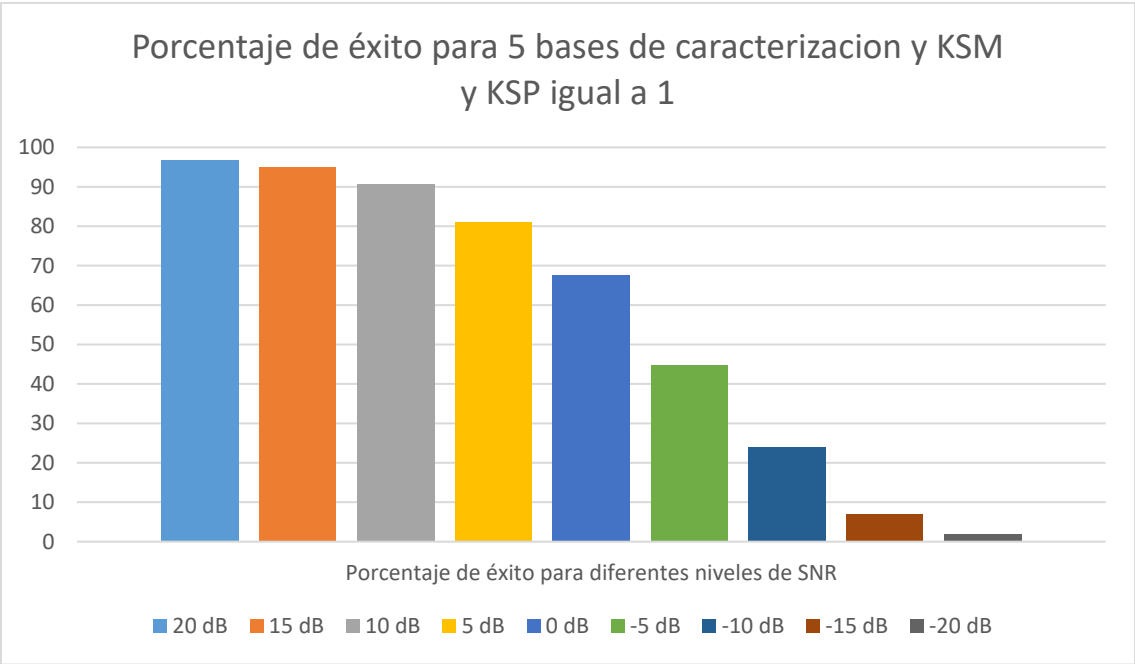


Figura 87 Gráfico de porcentaje de éxito en el reconocimiento de insectos con una configuración de 5 bases de caracterización para los insectos y los tipos de ruido y un valor 1 para sparseness y smoothness. El total de operaciones por cada nivel de SNR han sido 2024.

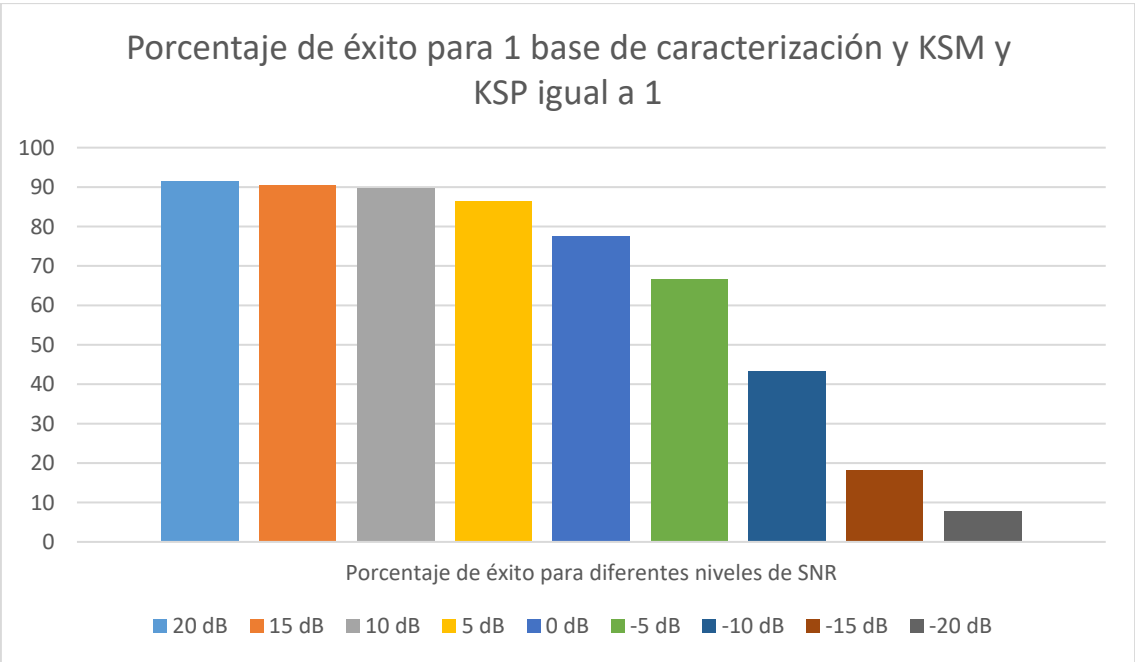


Figura 88 Gráfico de porcentaje de éxito en el reconocimiento de insectos con una configuración de 1 base de caracterización para los insectos y los tipos de ruido y un valor 1 para sparseness y smoothness. El total de operaciones por cada nivel de SNR han sido 2024.

En las gráficas superiores, observamos como para niveles positivos de SNR el comportamiento del sistema mejora al utilizar 3 bases de caracterización. Cuando comenzamos a disminuir el nivel de SNR por debajo de 0 dB, incluyendo éste, se observa una mejora en el reconocimiento al utilizar 1 base de caracterización. Pasando por ejemplo de un porcentaje de éxito en el reconocimiento del 47% en el caso de utilizar 3 bases a un 66% usando 1 sola base de caracterización.

Por tanto, la combinación más adecuada de parámetros para la identificación mediante esta técnica tiene dos soluciones posibles: 3 bases de caracterización de insecto y 3 bases para los diferentes tipos de ruido (blanco, impulsivo y armónico) para niveles elevados de SNR y a partir de 0 dB lo más adecuado es utilizar una sola base de caracterización tanto de ruido como de insecto. Para los valores de sparseness y smoothness hemos visto que el primer término no afecta a la identificación, por ello se han utilizado los mismos valores para ambos. Los valores más recomendados son 1 y 3, siendo el valor 1 el que más porcentaje de éxito produce.

Referente a los insectos 13 y 17, deben de utilizarse unos valores de caracterización iguales a los usados para valores negativos de SNR para lograr identificar, en muy pocas ocasiones y para valores muy elevados de SNR, el insecto estudiado.

Por tanto, para finalizar este Trabajo Fin de Grado, se indican las configuraciones más exitosas en el reconocimiento de insectos:

- General SNR superior a 0 dB
 - Número de bases de caracterización (insecto y ruidos): 3
 - Valor de sparseness y smoothness: 1
- General SNR igual o inferior a 0 dB
 - Número de bases de caracterización (insecto y ruidos): 1
 - Valor de sparseness y smoothness: 1
- Caso insecto 13 17 y 22
 - Número de bases de caracterización (insecto y ruidos): 1
 - Valor de sparseness y smoothness: 1

REFERENCIAS BIBLIOGRÁFICAS

- [1] CABAÑAS-MOLERO, P., et al. Compositional model for speech denoising based on source/filter speech representation and smoothness/sparseness noise constraints. *Speech Communication*, 2016, vol. 78, p. 84-99.
- [2] KIM, Hyunsoo; PARK, Haesun. Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. *Bioinformatics*, 2007, vol. 23, no 12, p. 1495-1502.
- [3] SCHMIDT, Mikkel N.; OLSSON, Rasmus Kongsgaard. Single-channel speech separation using sparse non-negative matrix factorization. En *Spoken Language Processing, ISCA International Conference on (INTERSPEECH)*. 2006.
- [4] SMARAGDIS, Paris; BROWN, Judith C. Non-negative matrix factorization for polyphonic music transcription. En *Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop on*. IEEE, 2003. p. 177-180.
- [5] DUONG, Viet-Hang, et al. Spatial dispersion constrained NMF for monaural source separation. En *Chinese Spoken Language Processing (ISCSLP), 2016 10th International Symposium on*. IEEE, 2016. p. 1-4.
- [6] HE, Guiqing, et al. Study on WorldView-2 Image Fusion Method Based on NMF and HCS Transform. En *Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2016 IEEE International Conference on*. IEEE, 2016. p. 548-553.
- [7] LEE, Daniel D.; SEUNG, H. Sebastian. Learning the parts of objects by non-negative matrix factorization. *Nature*, 1999, vol. 401, no 6755, p. 788-791.
- [8] PAATERO, Pentti; TAPPER, Unto. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 1994, vol. 5, no 2, p. 111-126.
- [9] MOORE, Holly. *MATLAB for Engineers*. Prentice Hall Press, 2014.

- [10] RODRÍGUEZ, Claudio; CERDA, Luis; PEREDO, Hernán. Detección de insectos causantes de daños en viveros de *Pinus radiata* de la Décima Región. *Bosque*, 1980, vol. 3, no 2, p. 73-76.
- [11] GARCÍA, José Lorenzo Meza; BÁEZ, Jesús Ricardo Camacho; MONDACA, Edgardo Cortes. Manejo integrado de plagas en granos almacenados. *Tecnologías de Granos y Semillas*, 1994, vol. 109.
- [12] PÉREZ, Nicolás, et al. Caracterización acústica de la colmena para la detección temprana de contaminación por pesticidas. *Cangüé*, 2014, vol. 35, p. 1-5.
- [13] CASALLASAS, Melanie Ramírez; YANINE, Mario Iván Ortiz; MOLINA, Jorge. Señales vibratorias en comportamiento de insectos: Producción, detección y métodos de estudio. *Sociedad Colombiana de Entomología-SOCOLEN*, 2013, p. 117.
- [14] GARRIGUES, José Vicente. ANALISIS ACÚSTICO DEL CANTO DE LA “CICADA ORNI”, VARIEDAD VALENCIANA.
- [15] CLEMENTE, María Eulalia; GARCÍA, María Dolores; PRESA, Juan José. *Clave de los géneros de saltamontes ibéricos (Orthoptera, Caelifera)*. EDITUM, 1987.
- [16] LEE, Daniel D.; SEUNG, H. Sebastian. Algorithms for non-negative matrix factorization. En *Advances in neural information processing systems*. 2001. p. 556-562.
- [17] Avisoft Bioacoustics © 2006-2015 [consulta: 1 febrero 2017]. Disponible en: <http://www.avisoft.com/sounds.htm#insects>
- [18] HU, Yi; LOIZOU, Philipos C. Subjective comparison and evaluation of speech enhancement algorithms. *Speech communication*, 2007, vol. 49, no 7, p. 588-601.
- [19] MARÍN-HENAO, María Isabel; BORRERO-DOMÍNGUEZ, Jeyson Javier; GONZÁLEZ-ALVARADO, Iván Camilo. Identificador y clasificador de insectos en campos de cultivo. Identifier and classifier of insects in field crops.